CSE 559A: Computer Vision

[credit: danjodon.deviantart.com]

Fall 2017: T-R: 11:30-1pm @ Lopata 101

Instructor: Ayan Chakrabarti (ayan@wustl.edu).
TAs: Abby Stylianou (abby@wustl.edu), Jarett Gross (jarett@wustl.edu)

http://www.cse.wustl.edu/~ayan/courses/cse559a/

Aug 29, 2017

INTRODUCTION

What is Computer Vision ?

Endow machines with the ability to make *sense*

of the physical world by looking at

images and videos

INTRODUCTION

What is Computer Vision ?

Endow machines with the ability to make *sense*

of the physical world by looking at

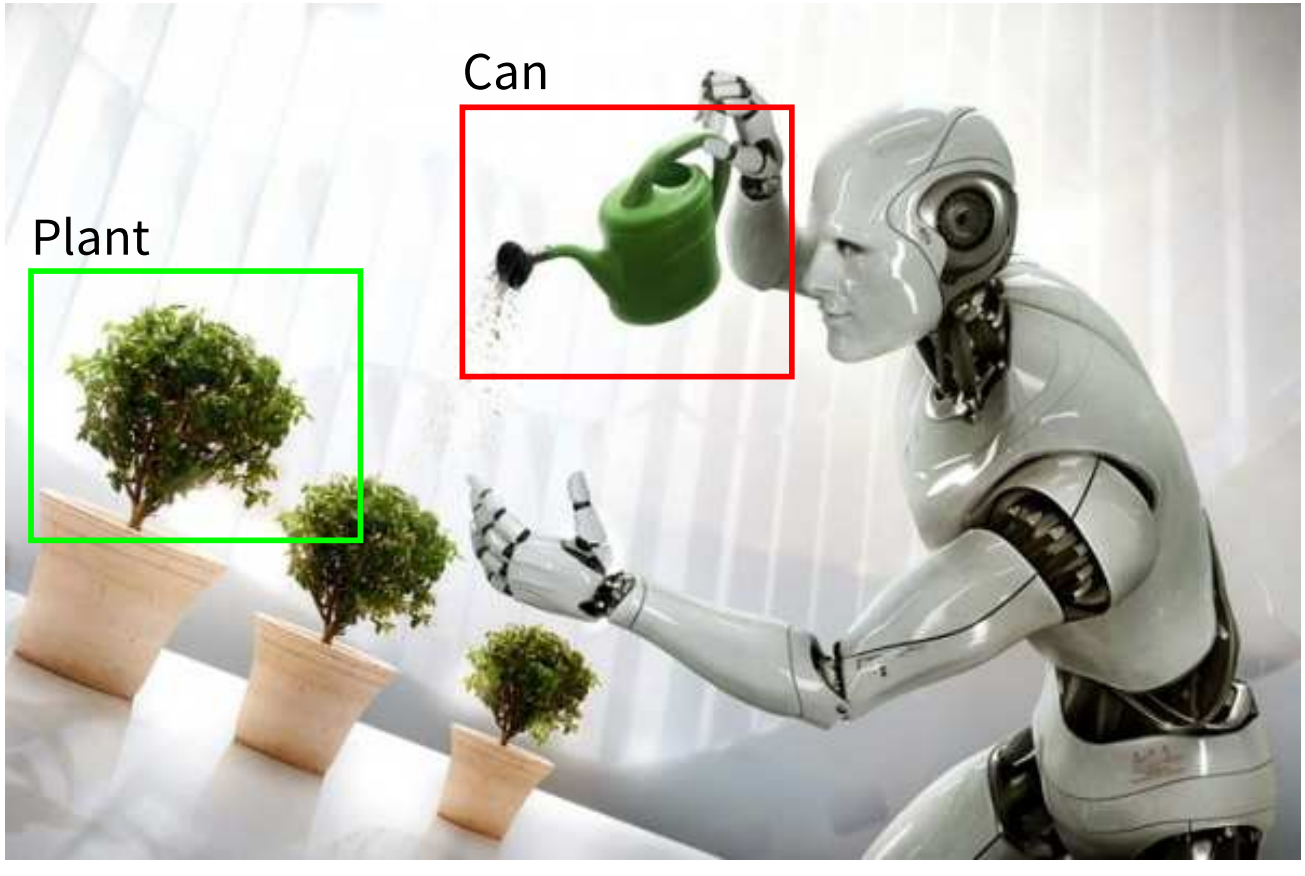
measurements of reflected light

INTRODUCTION

[credit: http://www.blutsbrueder-design.com]

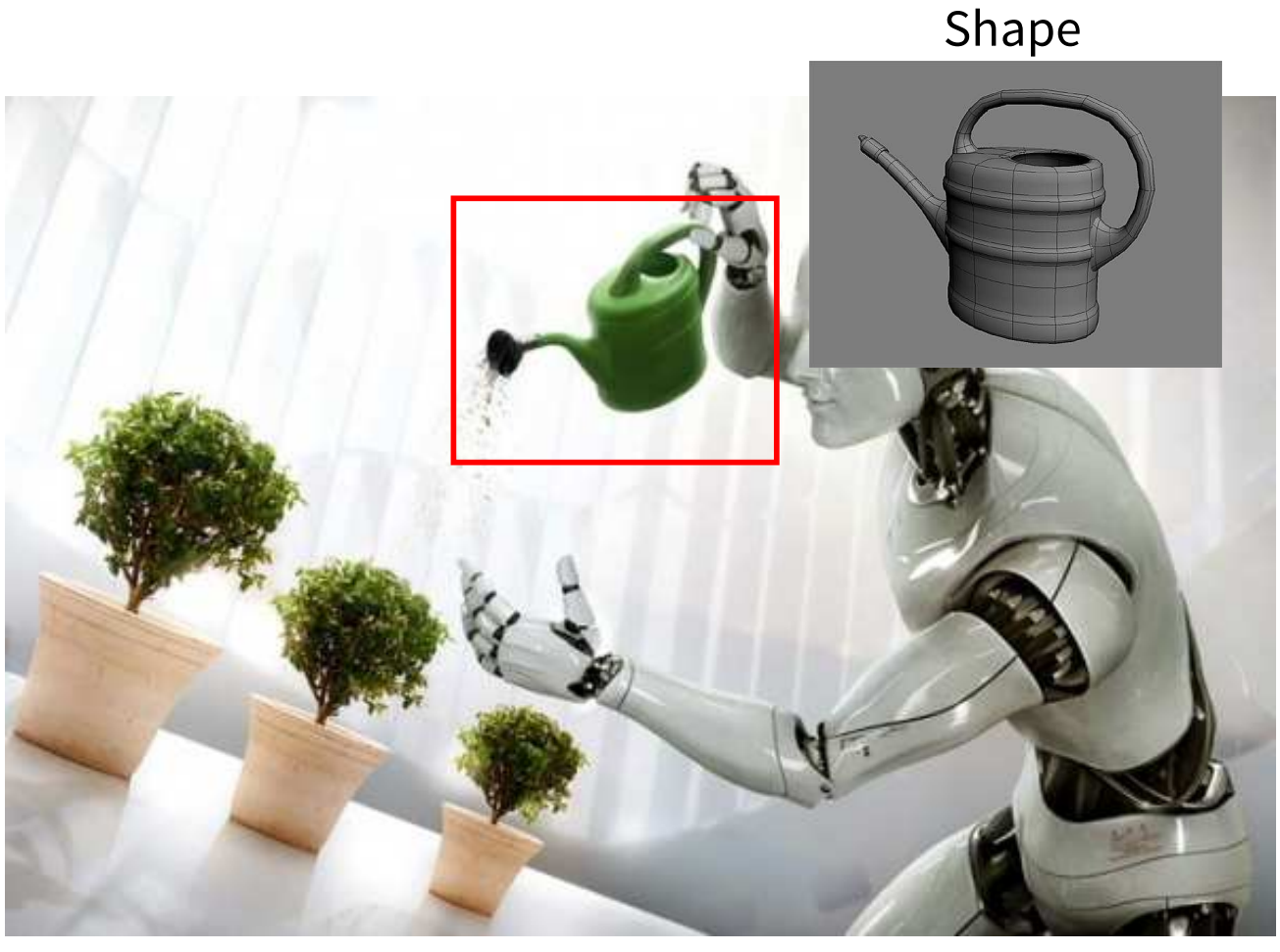
INTRODUCTION

INTRODUCTION



Recognize Objects

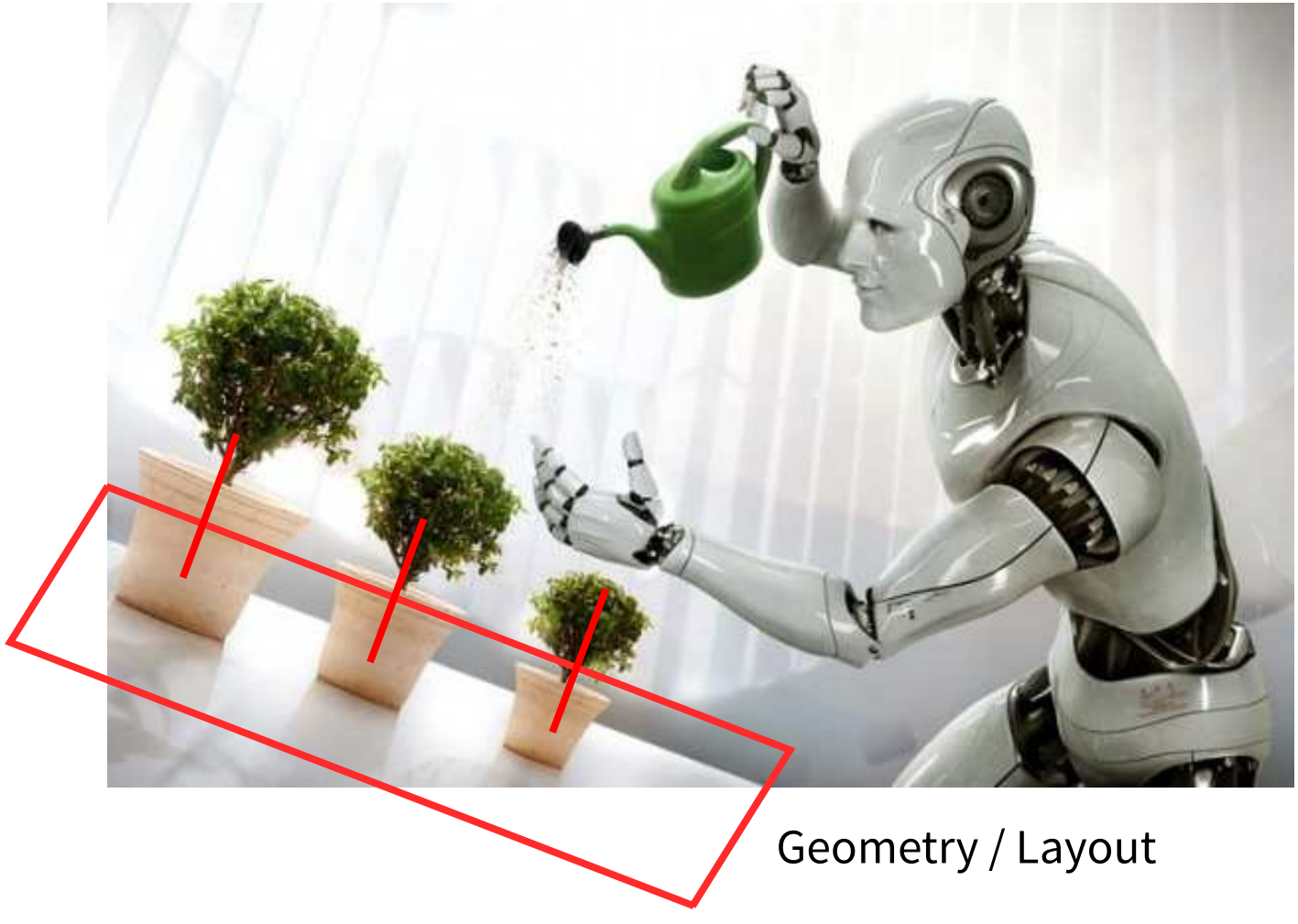
INTRODUCTION



INTRODUCTION

Classify Scene

INTRODUCTION

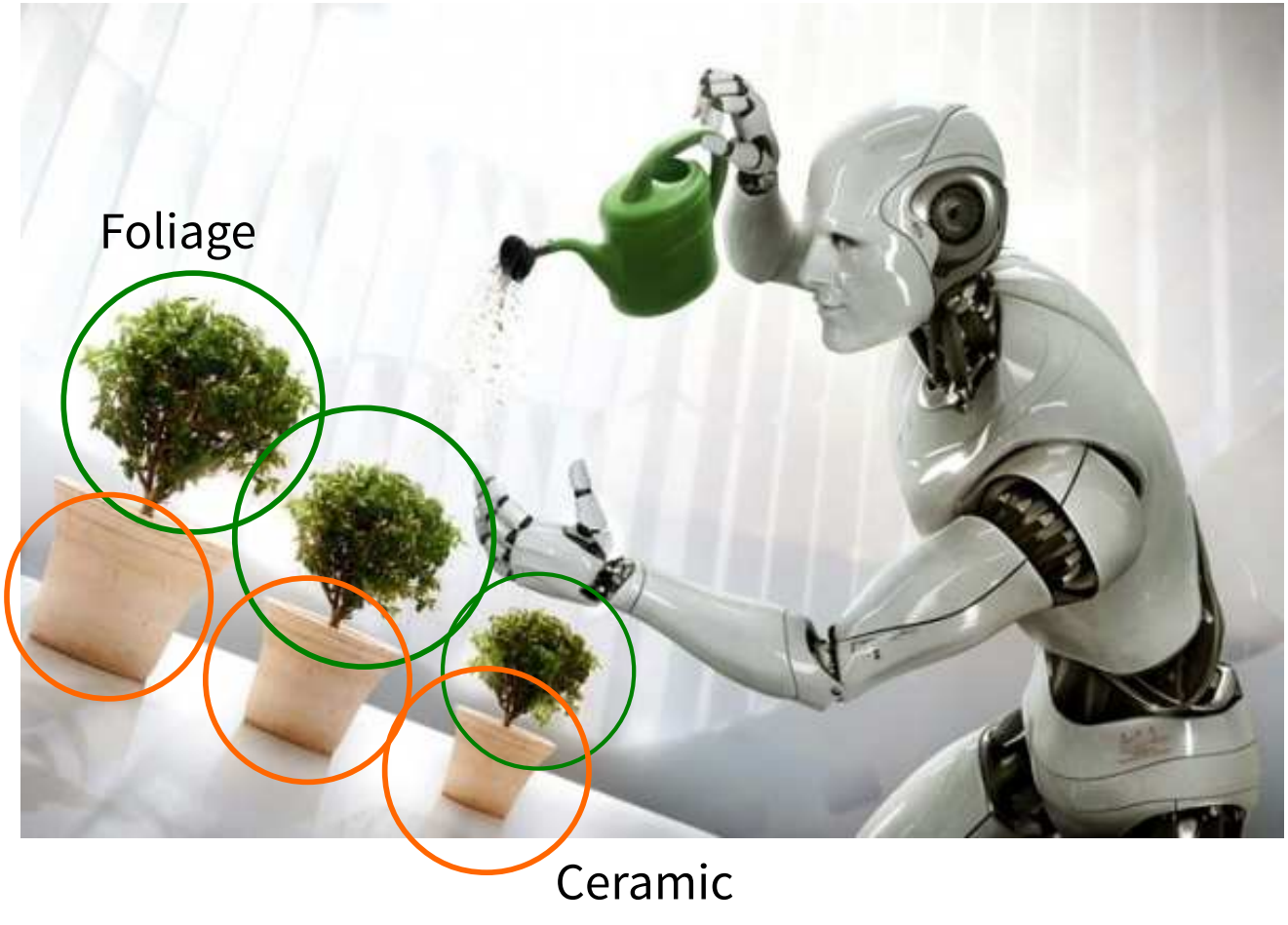


Geometry / Layout

INTRODUCTION

INTRODUCTION

Identify Materials

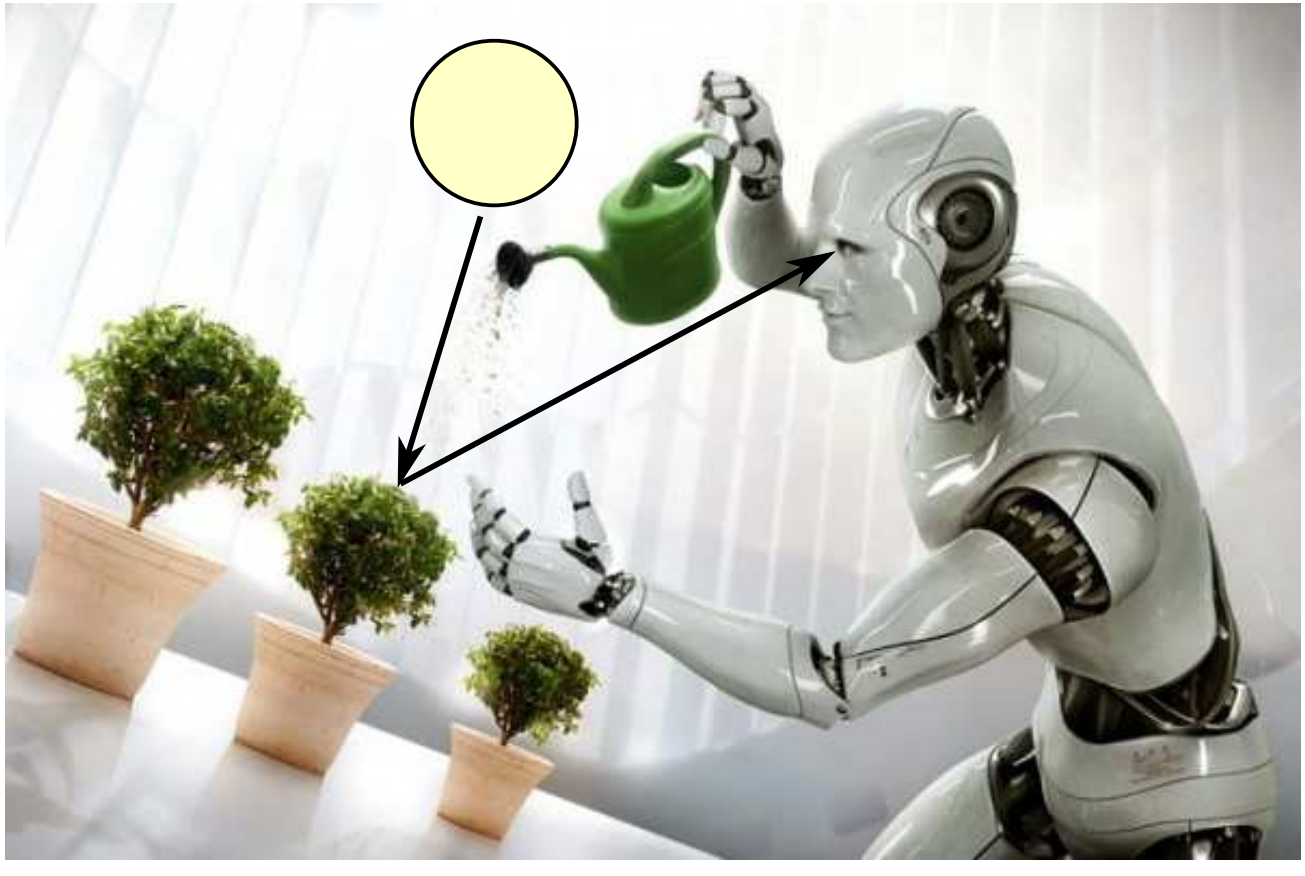


INTRODUCTION

Surface Properties

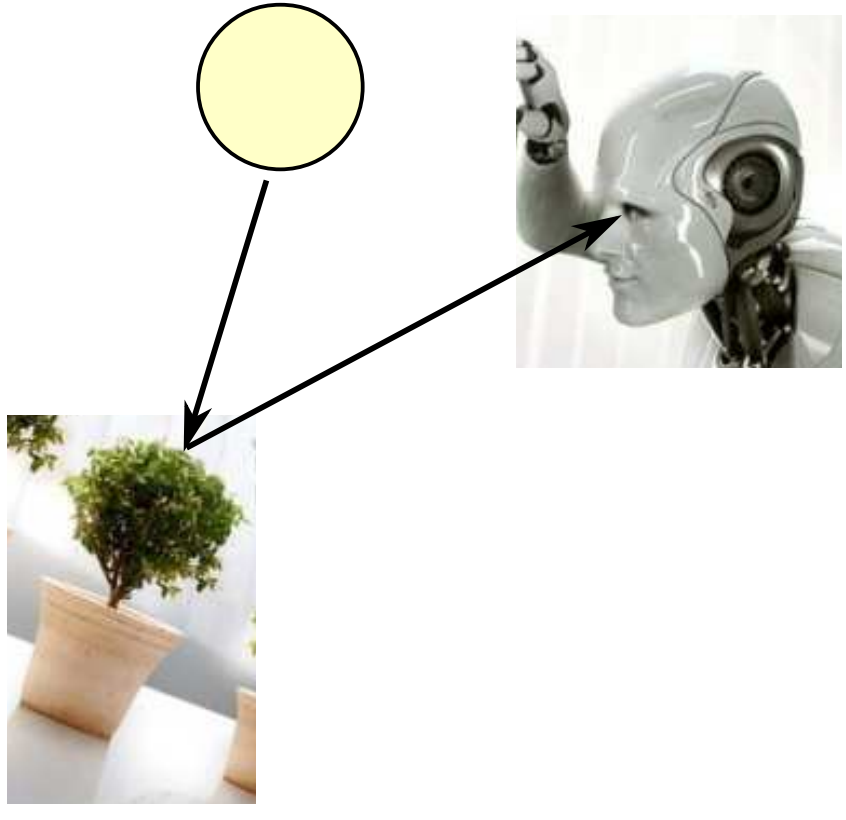
WHY IS THIS HARD ?

WHY IS THIS HARD ?



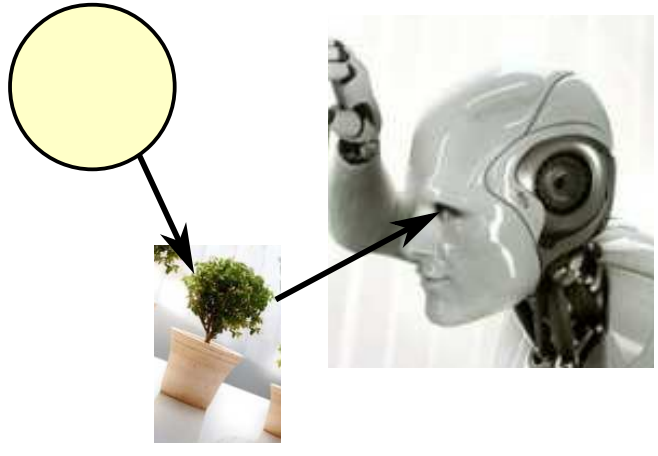
WHY IS THIS HARD ?

Scale Ambiguity



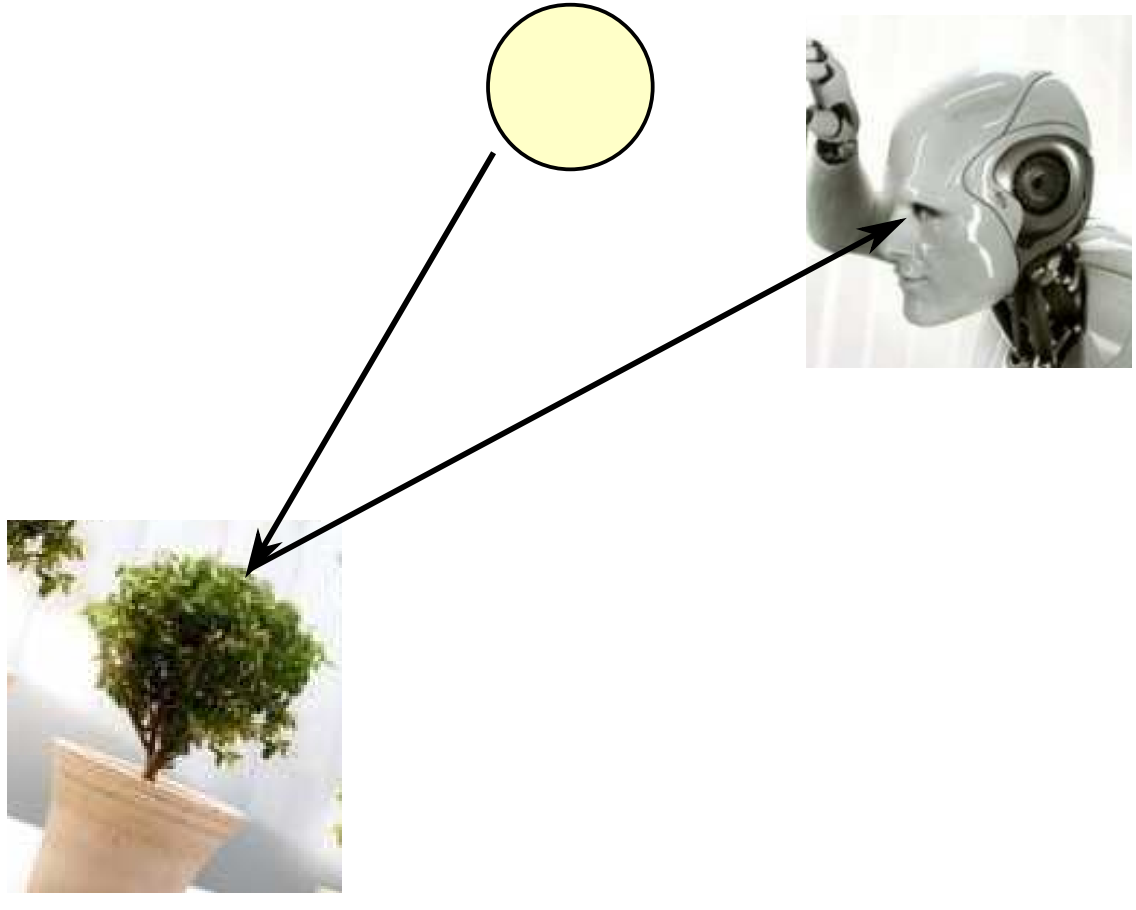
WHY IS THIS HARD ?

Scale Ambiguity



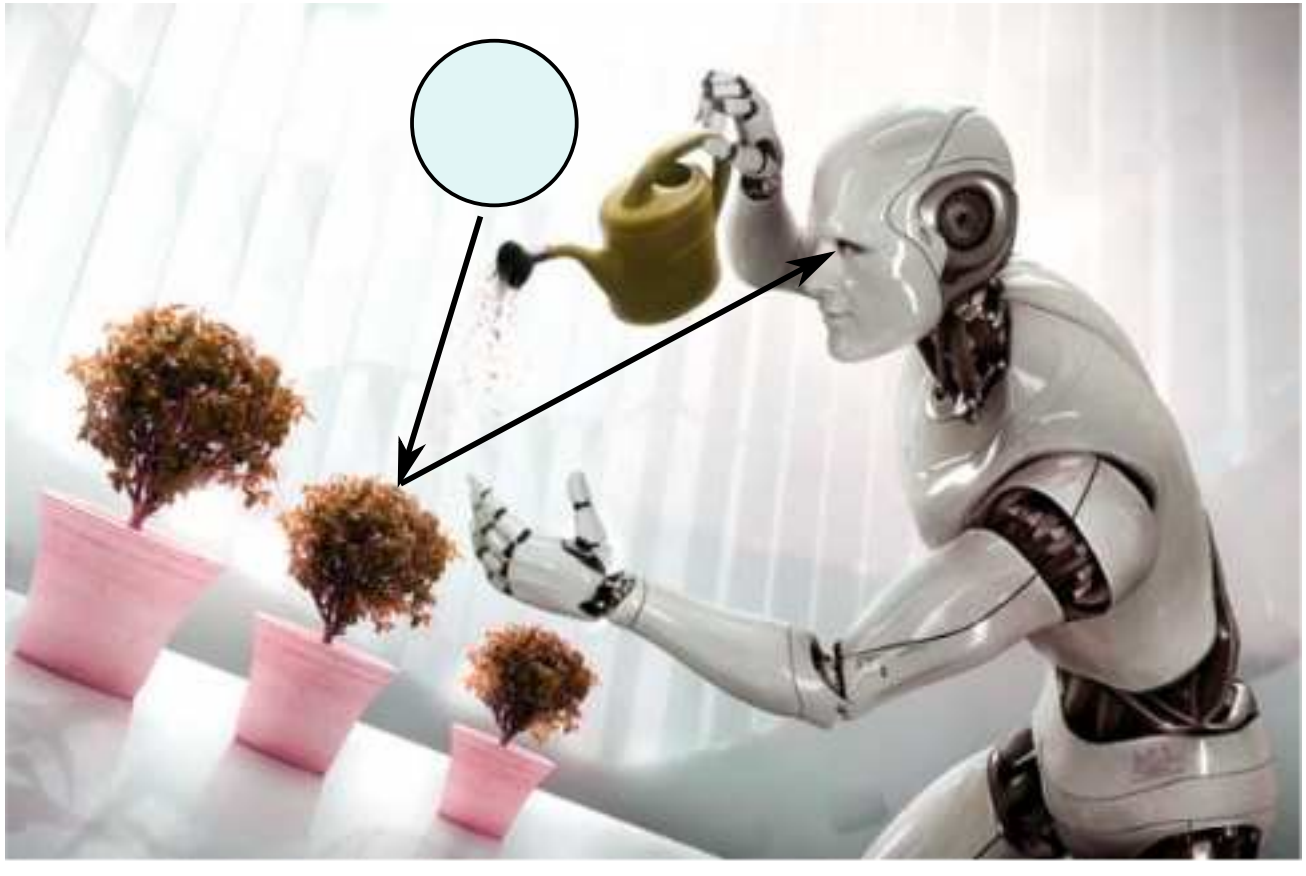
WHY IS THIS HARD ?

Scale Ambiguity



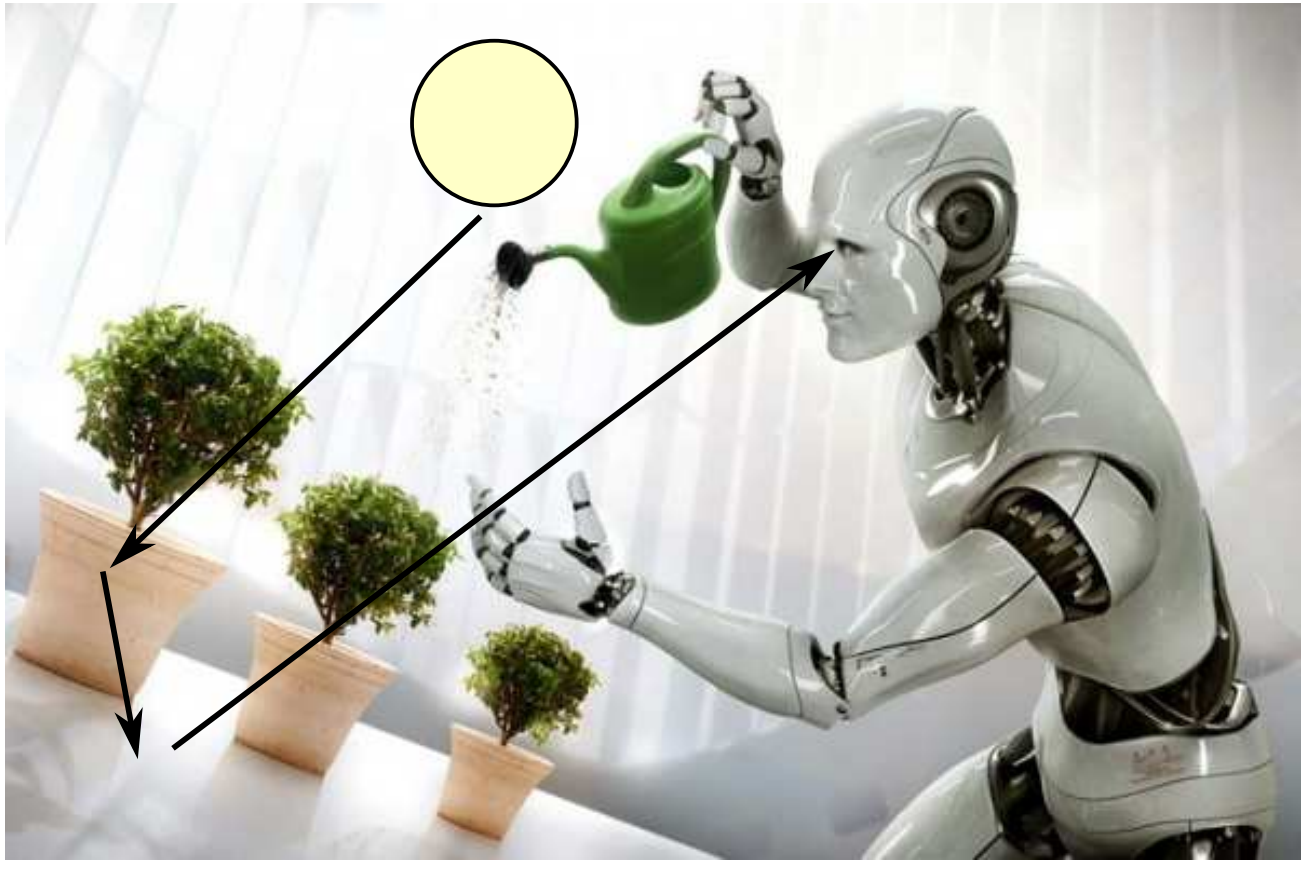
WHY IS THIS HARD ?

Illuminant / Surface Color Ambiguity



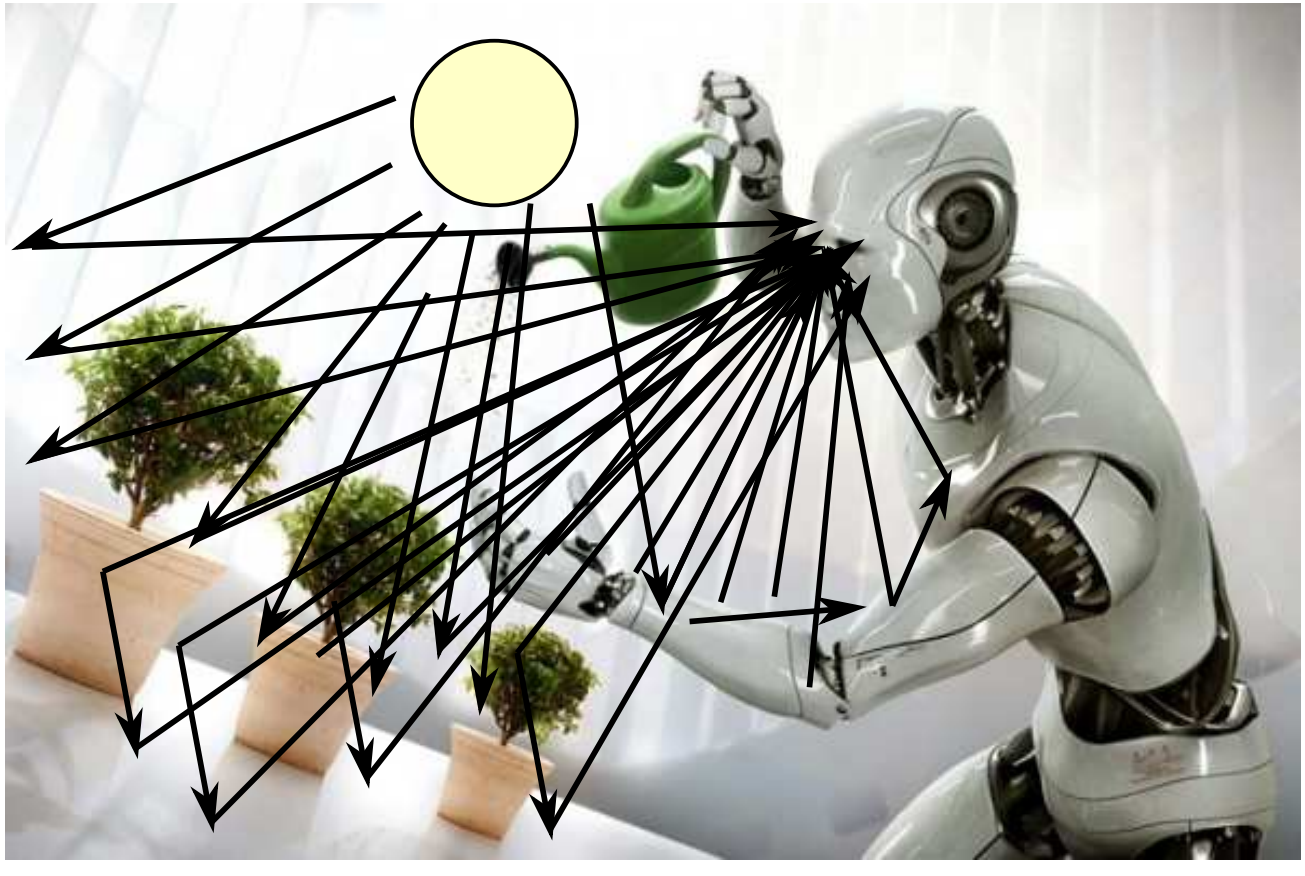
WHY IS THIS HARD ?

Indirect Reflection



WHY IS THIS HARD ?

Observe Super-position of Multiple Paths



WHY IS THIS HARD ?

Variation in Appearance

WHY IS THIS HARD ?

Variation in Appearance

WHY IS THIS HARD ?

make sense of the physical world

from measurements of reflected light

Seems hopeless .... except that
humans, animals, birds, insects are able to do it

SO HOW DOES IT WORK ?

By making assumptions about and exploiting structure in the natural world

SO HOW DOES IT WORK ?

BROAD OVERVIEW OF (MANY A) VISION ALGORITHM

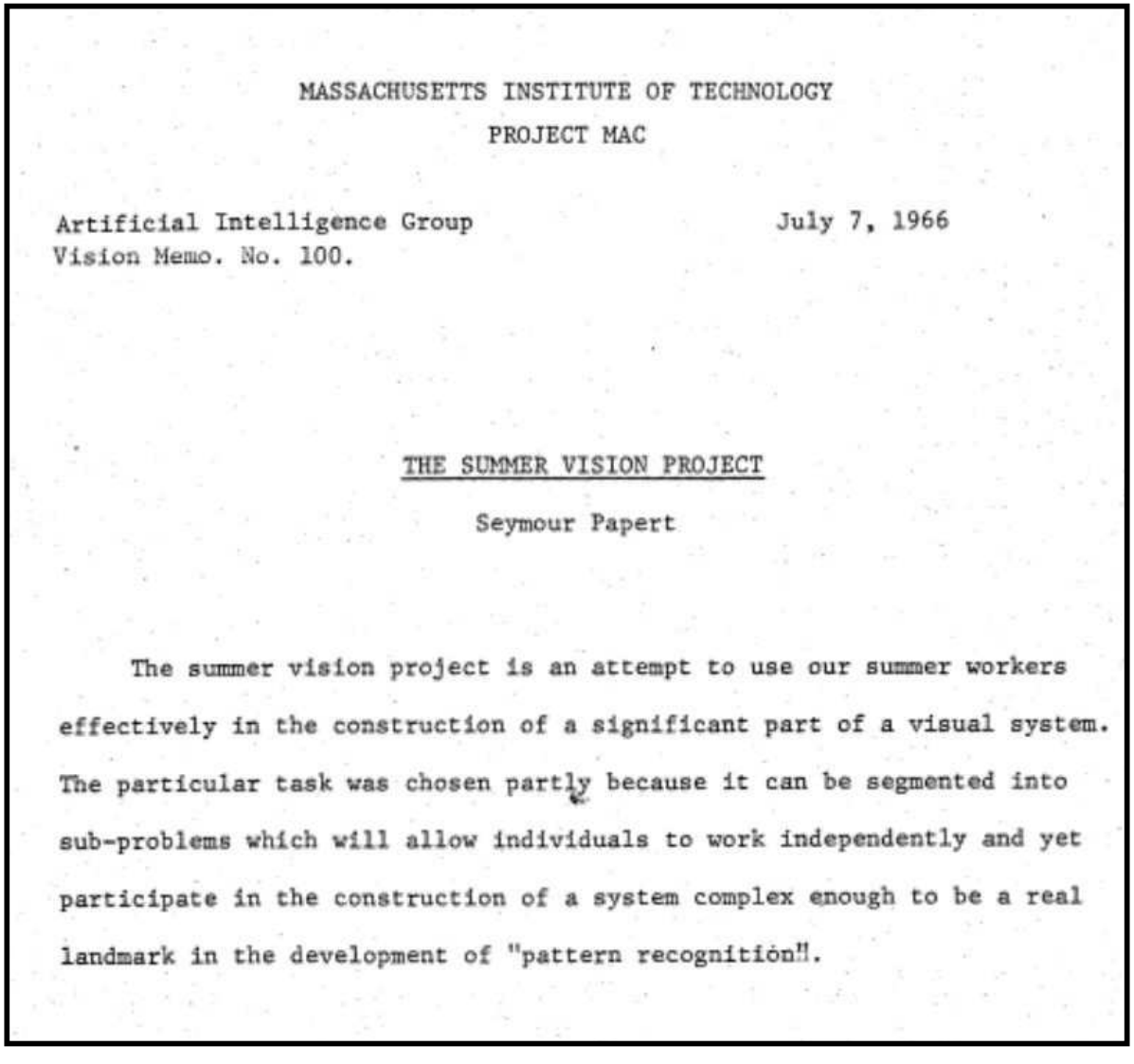
- Build physical model for mapping from scene to image: $I = f(S)$
- Learn a statistical model $p(S)$ for what scene values are likely.
- Given an image, consider the *feasible set* of scenes: $\mathcal{F}^{-1}(I) = \{S : I = f(S)\}$
- Choose the most likely answer from the feasible set: $\hat{S} = \underset{S \in \mathcal{F}^{-1}(I)}{\arg\max}\ p(S)$

ILLUSIONS

Deliberate / Artistic

Natural (rare!)

HISTORY OF CV



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Attempt 1:

Let's finish this over the summer.

Slide via Lana Lazebnik

HISTORY OF CV

"Simple"
Pattern Recognition

Slide via Lana Lazebnik

HISTORY OF CV

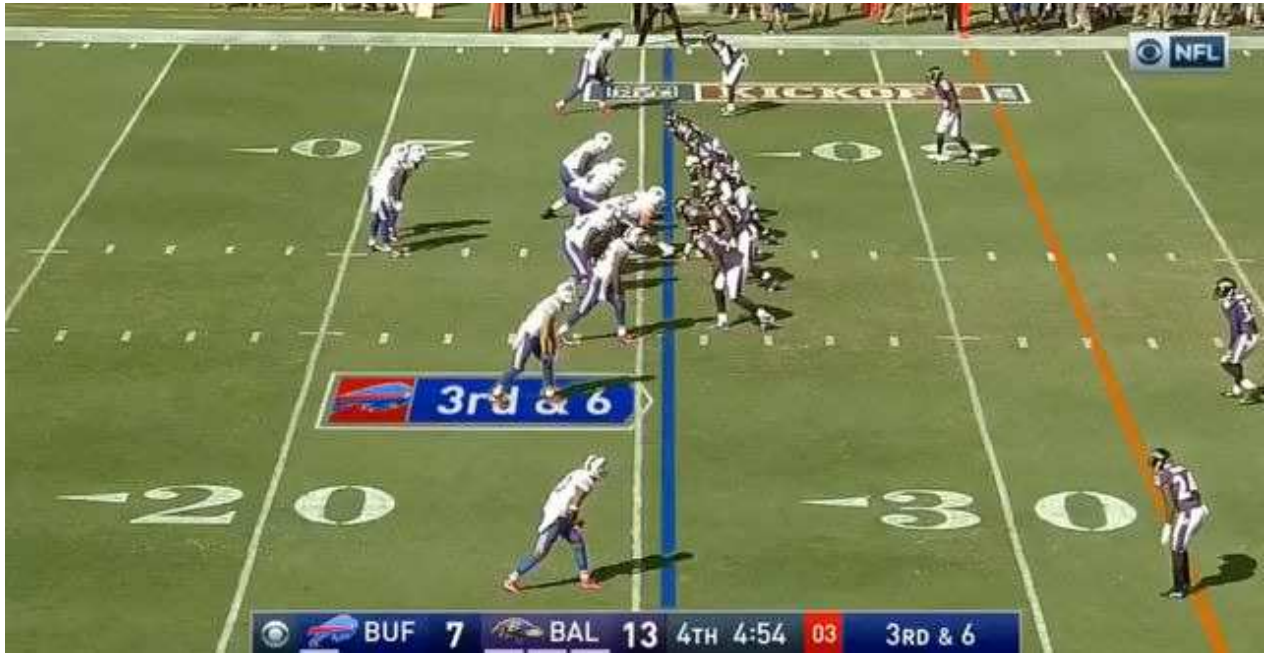
Face Detection
in Cameras

HISTORY OF CV

Smile Detection in Cameras

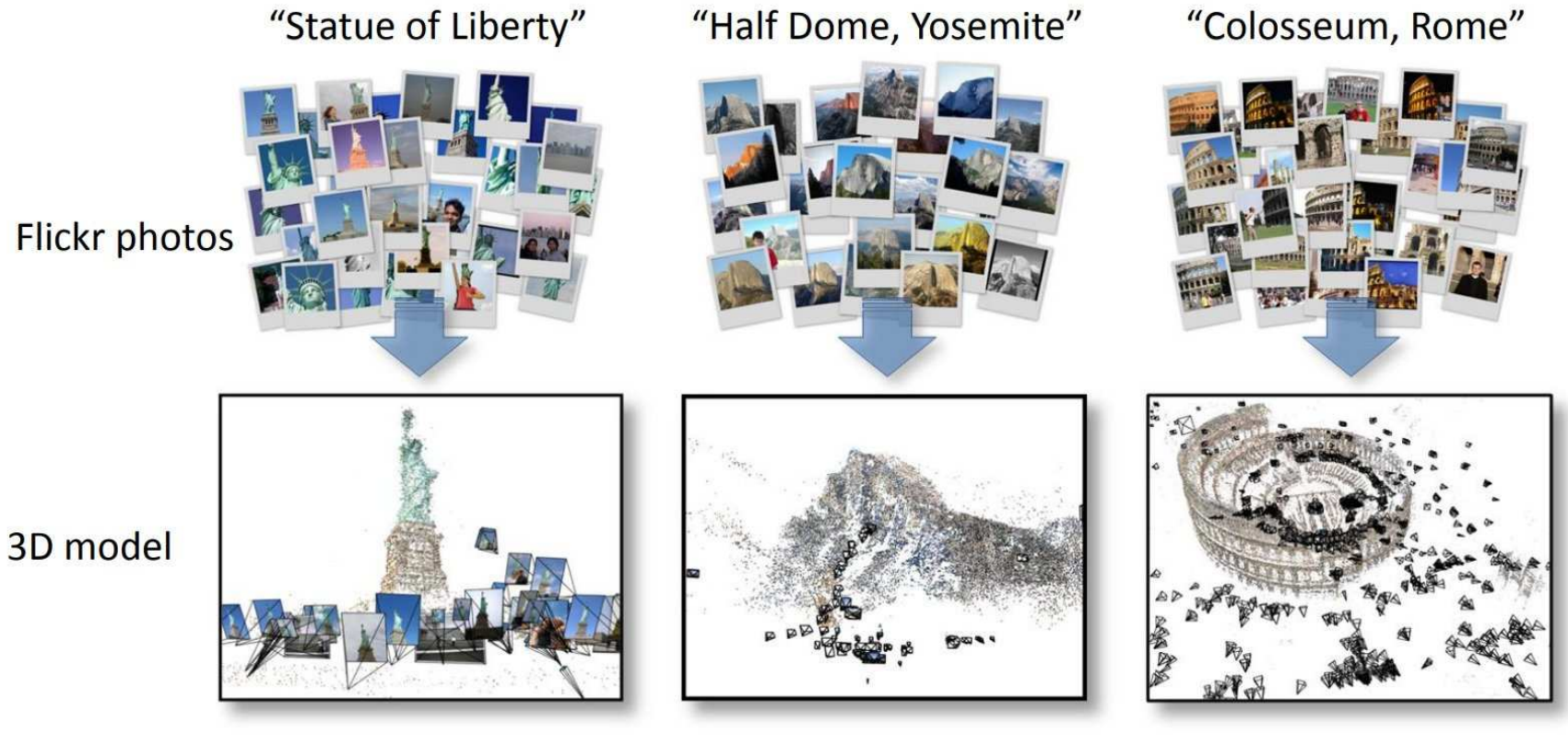
Slide via Lana Lazebnik

HISTORY OF CV



Elementary
Augmented Reality

HISTORY OF CV



Noah Snavely, Sameer Agarwal, Ian Simon, Steve Seitz,
Richard Szeliski, Yasutaka Furukawa, Brian Curless

Slide via Noah Snavely

HISTORY OF CV

Face2Face: Real-time Face Capture and Reenactment of RGB Videos

Justus Thies [1] Michael Zollhöfer [2] Marc Stamminger [1] Christian Theobalt [2] Matthias Nießner [3]

[1] University of Erlangen-Nuremberg [2] Max Planck Institute for Informatics [3] Stanford University

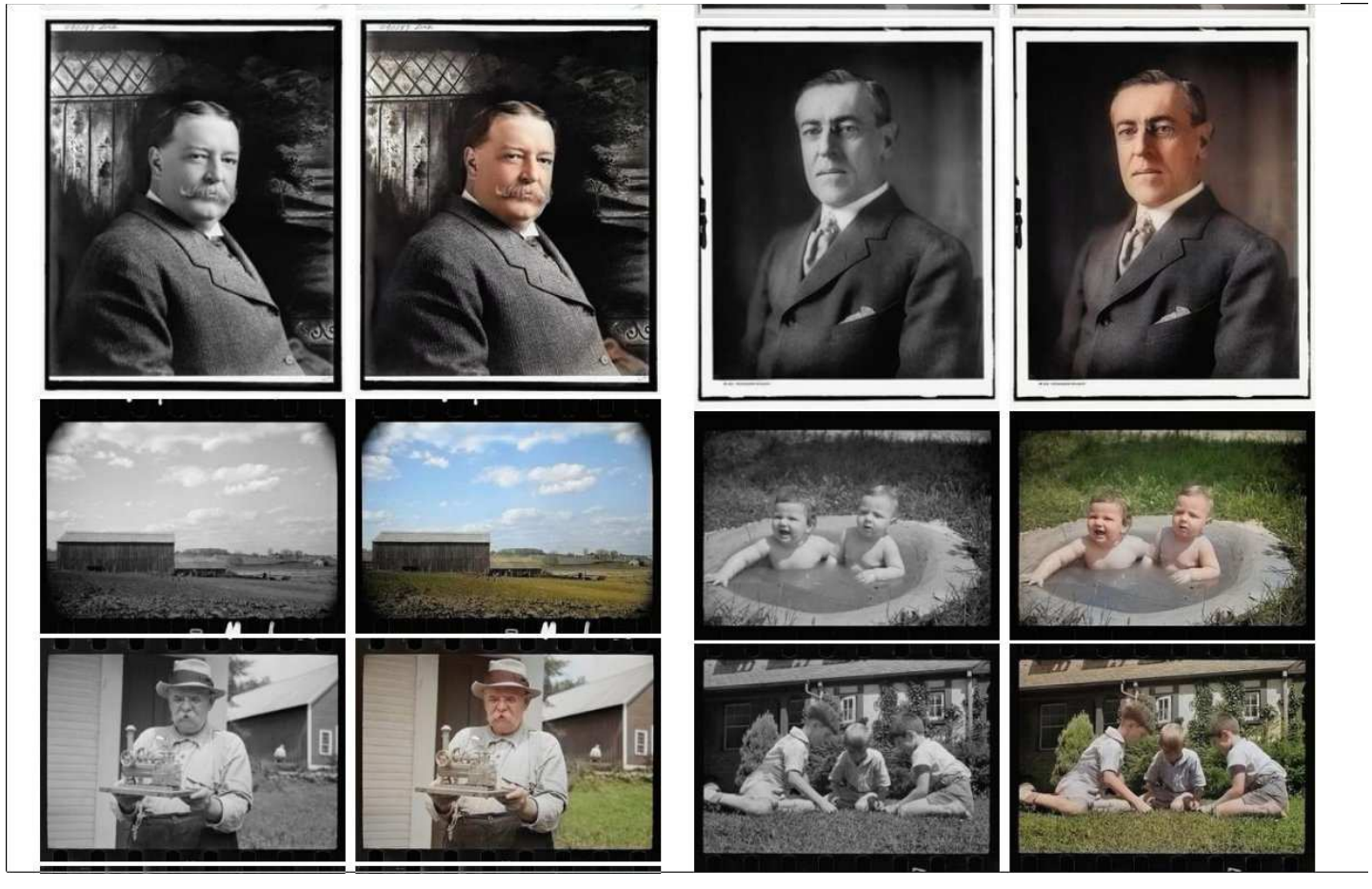
Proc. Computer Vision and Pattern Recognition (CVPR), IEEE, June 2016

HISTORY OF CV

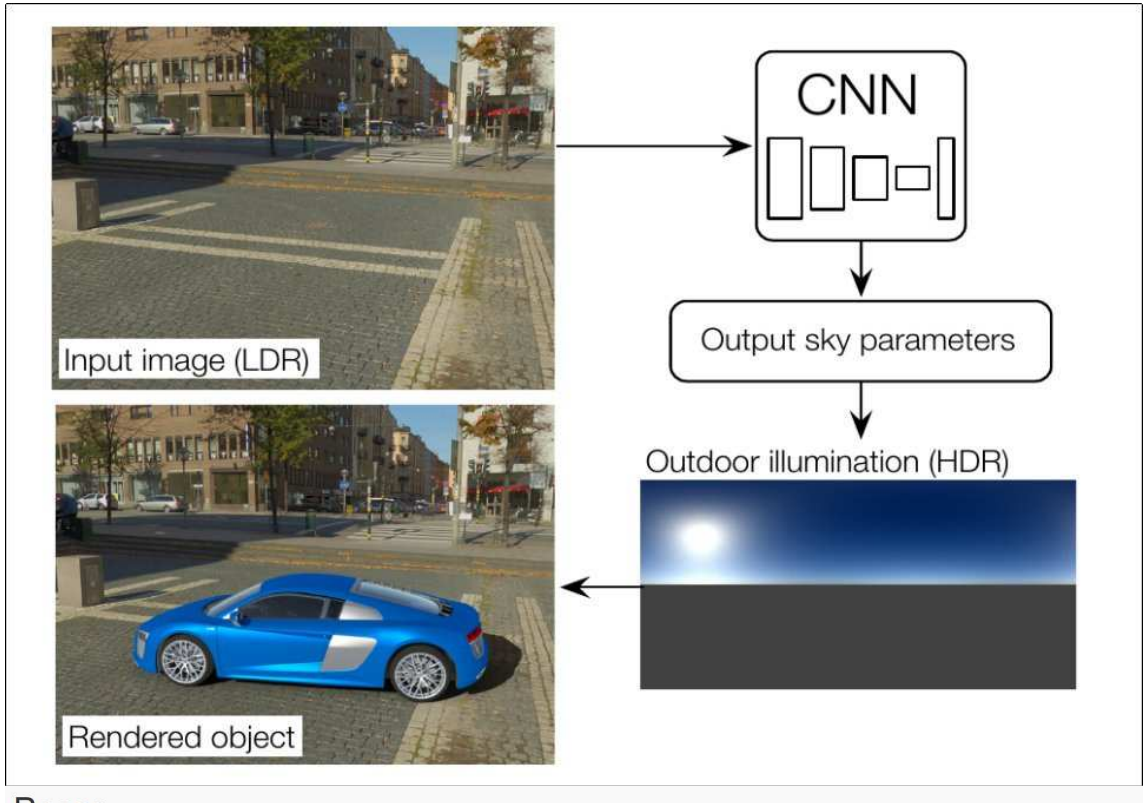
Learning Representations for Automatic Colorization

Gustav Larsson Michael Maire Gregory Shakhnarovich

ECCV 2016



HISTORY OF CV



Paper

Yannick Hold-Geoffroy, Kalyan Sunkavalli, Sunil Hadap, Emiliano Gambaretto and Jean-François Lalonde
Deep Outdoor Illumination Estimation
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
[arXiv pre-print] [BibTeX]

HISTORY OF CV

Scanning and Printing a 3D Portrait of President Barack Obama

Smithsonian Digitization Program Office and *USC Institute for Creative Technologies*

From Paul Debevec and Colleagues

HISTORY OF CV

Image-to-Image Translation with Conditional Adversarial Nets

Phillip Isola Jun-Yan Zhu Tinghui Zhou Alexei A. Efros

Univerity of California, Berkeley
In CVPR 2017

HISTORY OF CV

TECHNOLOGY

The New York Times

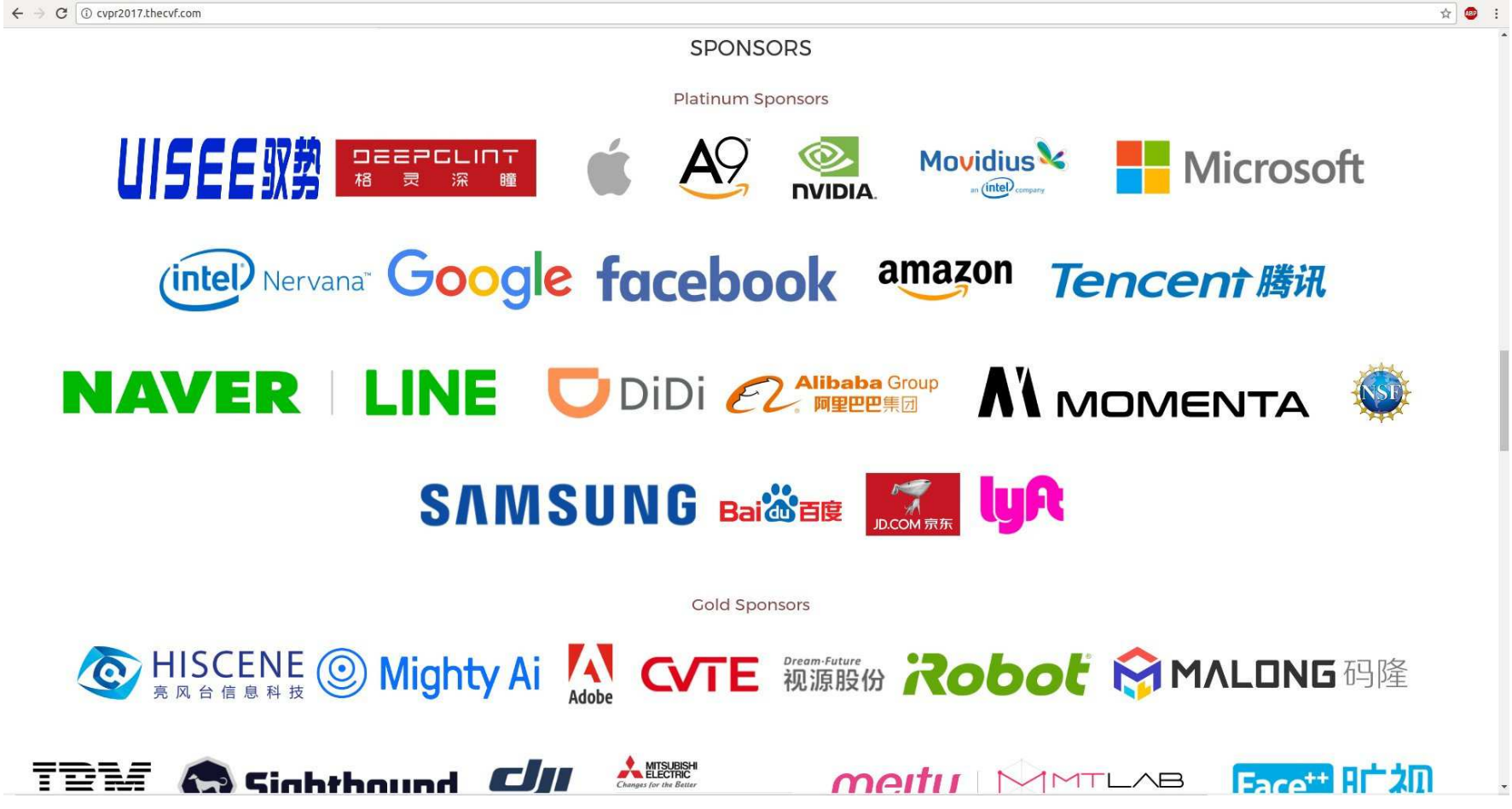
For Now, Self-Driving Cars Still Need Humans

By JOHN MARKOFF JAN. 17, 2016

Slide via Lana Lazebnik

HISTORY OF CV

Sponsor roster @ CVPR 2017



HISTORY OF CV

Sponsor roster @ CVPR 2017

HISTORY OF CV

Sponsor roster @ CVPR 2017

HISTORY OF CV

Sponsor roster @ CVPR 2017

HISTORY OF CV

Lots of Exciting Research & Potential for Real World Impact !

THIS COURSE

All information @ **http://www.cse.wustl.edu/~ayan/courses/cse559a/**

SYLLABUS OVERVIEW

- Image formation, representation, and processing
- Low level vision
  - Photometric
  - Geometric
  - Motion
- Segmentation & Grouping
- High Level Vision (reasons with "semantic" knowledge)
  - Using deep convolutional neural networks

Slides will be posted on course website after class

THIS COURSE

All information @ **http://www.cse.wustl.edu/~ayan/courses/cse559a/**

PREREQUISITES

- Programming: problem sets will be in Python.
- Math: Linear Algebra, Vector Calculus, Probability and Statistics.
- Also helpful: Prior coursework on machine learning, signal processing, graphics.

Please fill out survey on course website (also sent to you by e-mail) ASAP

THIS COURSE

All information @ **http://www.cse.wustl.edu/~ayan/courses/cse559a/**

- **5 Problem Sets**
  - Math (answers to be typeset in LaTeX) and Programming (in Python)
  - 15% x 5 = 75% of your grade
  - To be done individually
  - Roughly every two weeks (see website)
  - **READ** collaboration and late policy
- **Final Project**
  - To be done individually
  - Open ended, mini research/implementation project
  - Choose topic (will post suggestions), submit brief proposal, and get feedback from us
  - End of Term: Report + 5 min presentation = 25% of your grade

NO EXAMS

THIS COURSE

All information @ **http://www.cse.wustl.edu/~ayan/courses/cse559a/**

- **Extra Credit: Class Participation = 5%**

- Sign up for Piazza !
  - Ask questions on piazza: more efficient than e-mail, can also be anonymous
  - Answer others' questions (subject to collaboration policy): counts towards credit

- Office Hours
  - Prof. Chakrabarti: Wed 9:30am-10:30am, Jolley Hall @ 205
  - TAs: TBD (will post poll on Piazza)
  - Do come and ask questions sooner rather than later
  - No participation credit for office hours

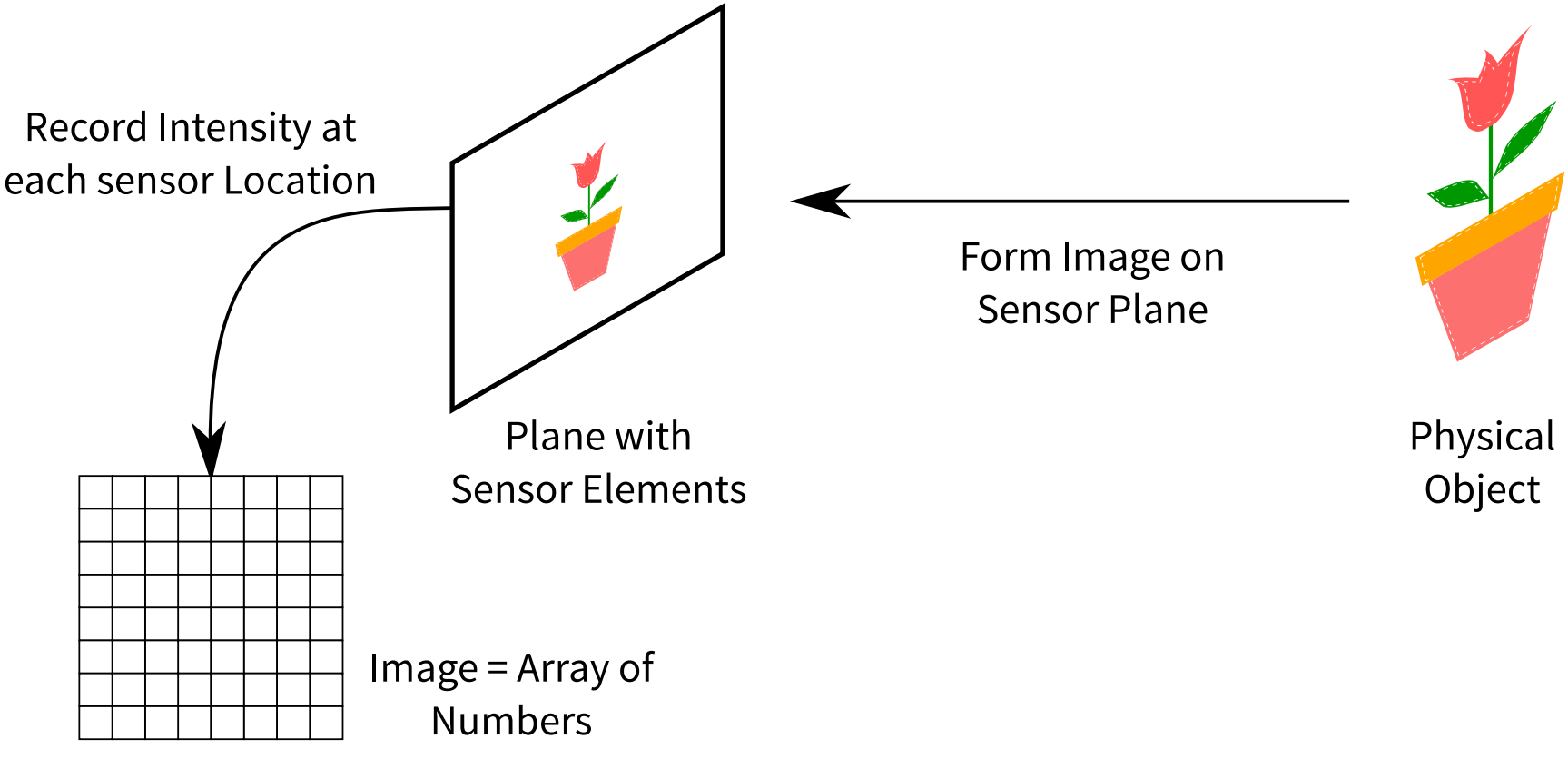
THIS COURSE

All information @ **http://www.cse.wustl.edu/~ayan/courses/cse559a/**

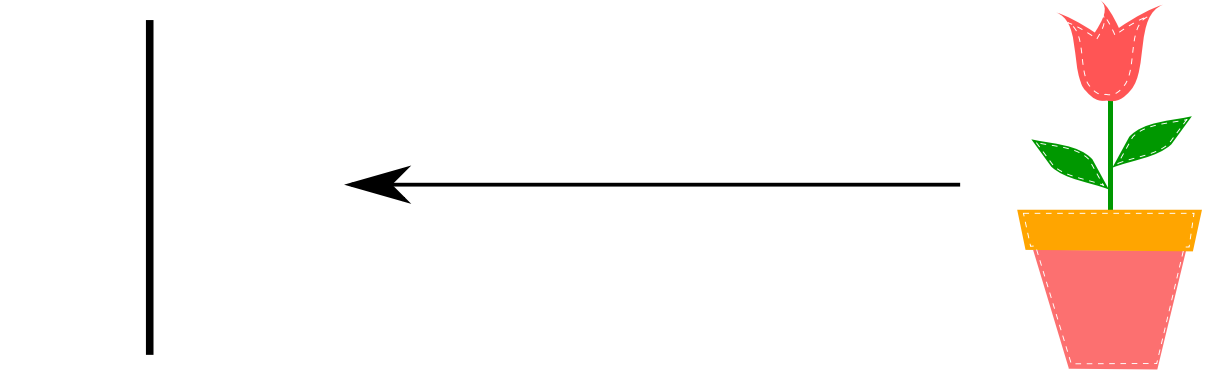
Questions ?

- Read the syllabus and the late and collaboration policies on the course website.
- Join Piazza
- Fill out survey

THE PINHOLE CAMERA

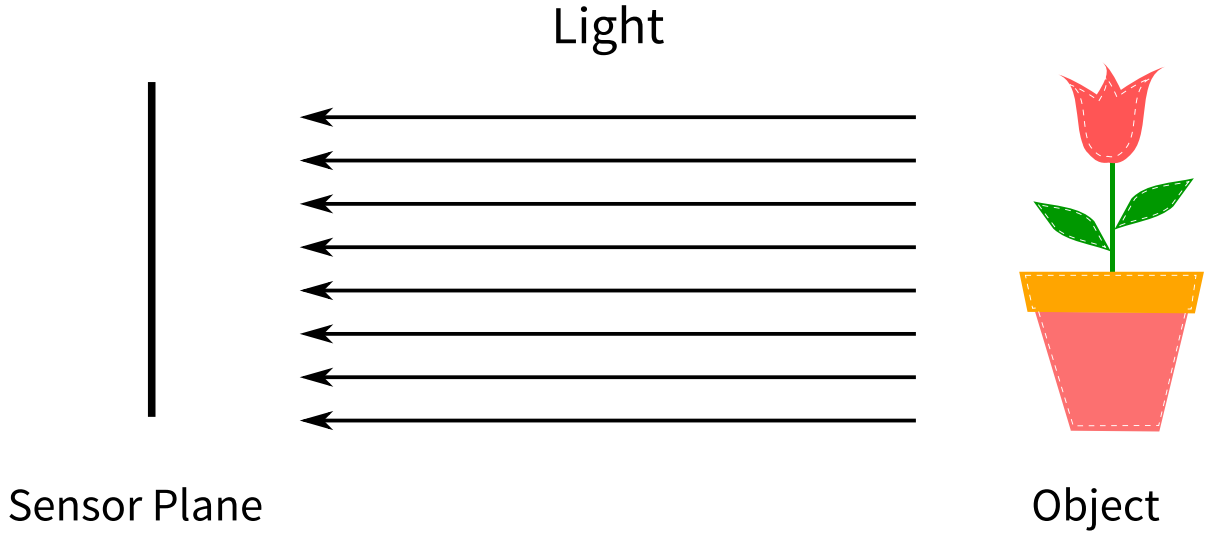


THE PINHOLE CAMERA



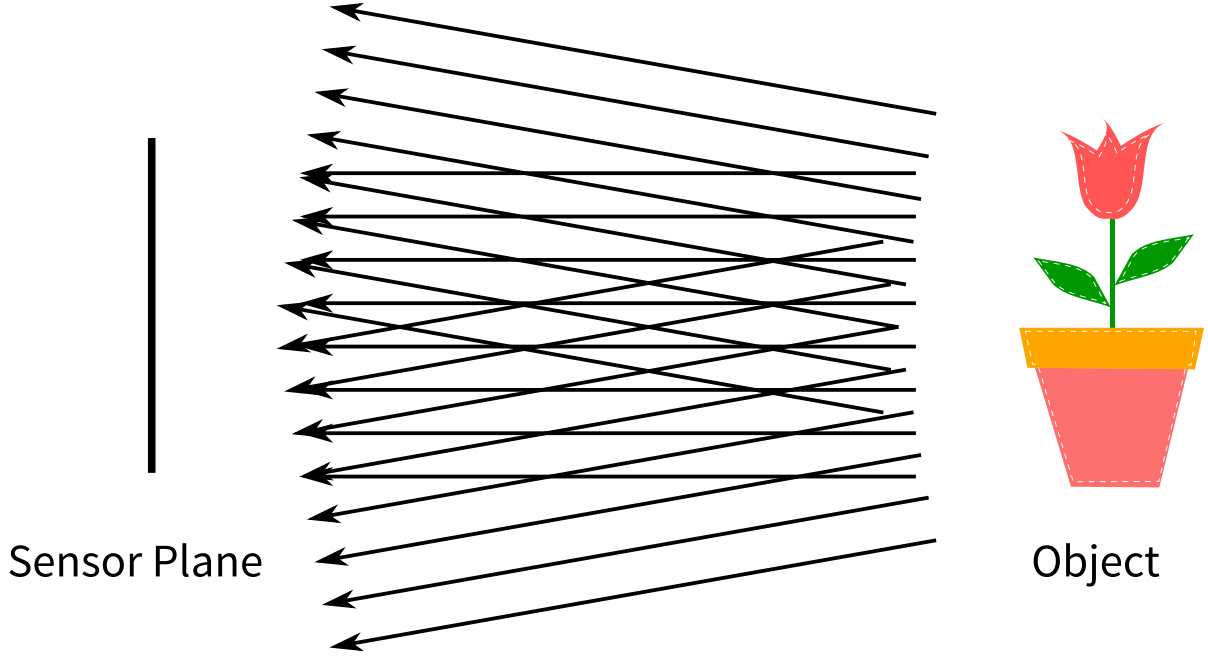
1D for simplicity

THE PINHOLE CAMERA



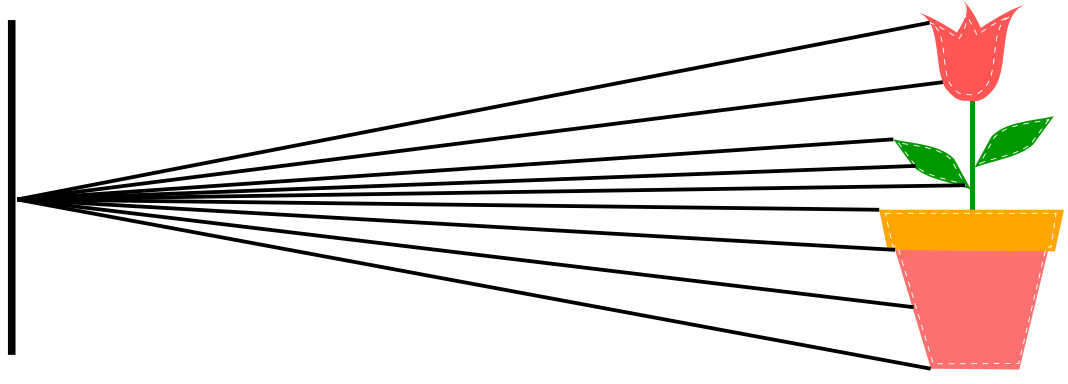
Why can't we get an image to form by just holding the sensor infront of the object ?

THE PINHOLE CAMERA



Because the object isn't just reflecting in straight lines!

THE PINHOLE CAMERA

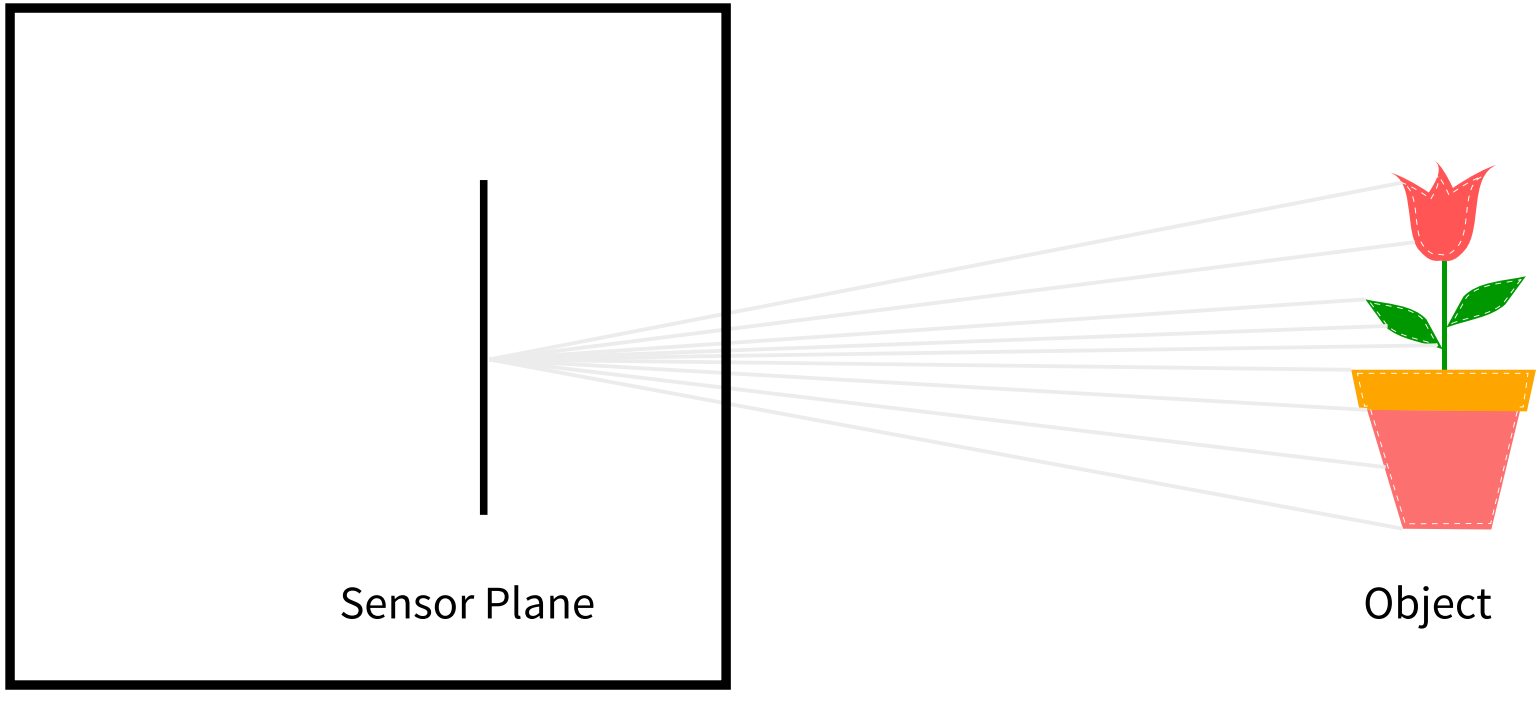


Sensor Plane

Object

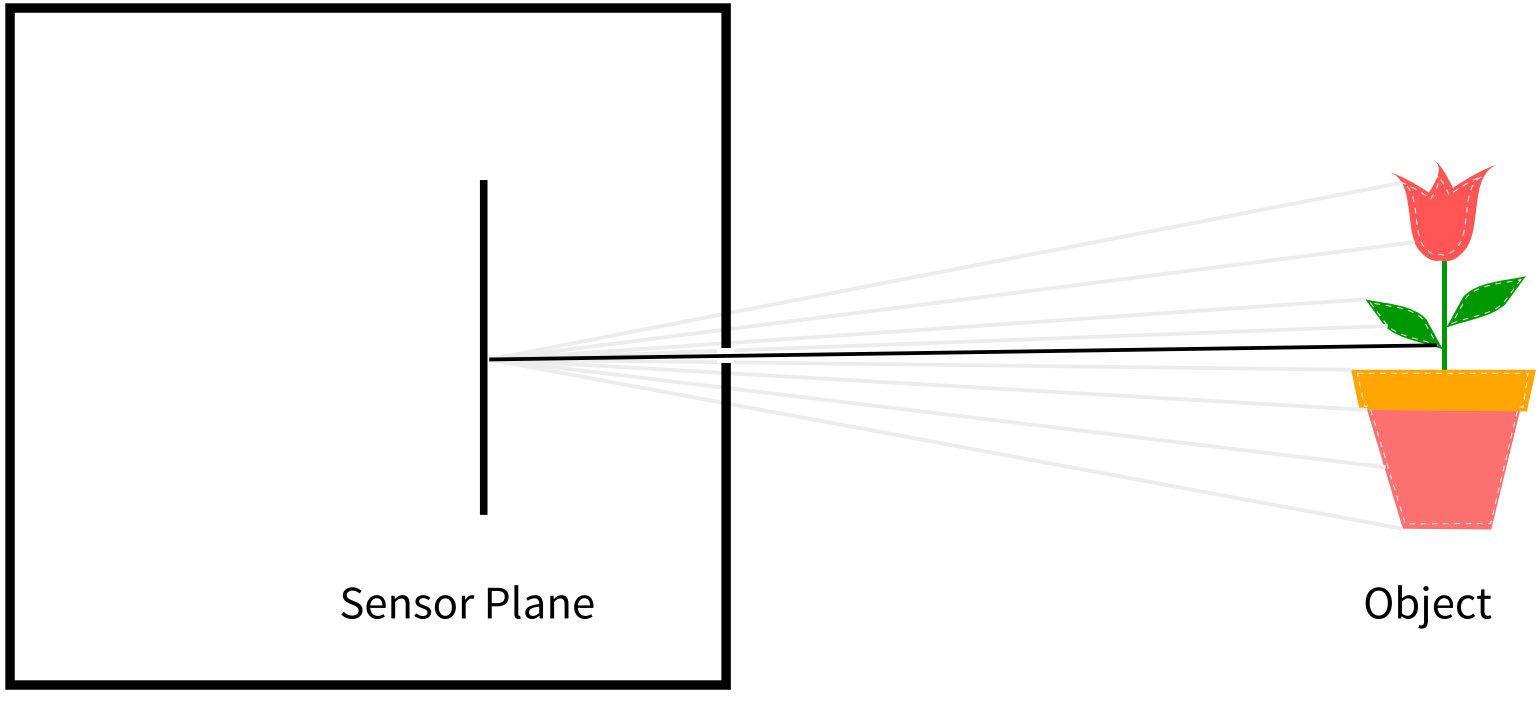
A point on sensor plane receives light
from many distinct points on the object.
No image since intensities "washed out".

THE PINHOLE CAMERA



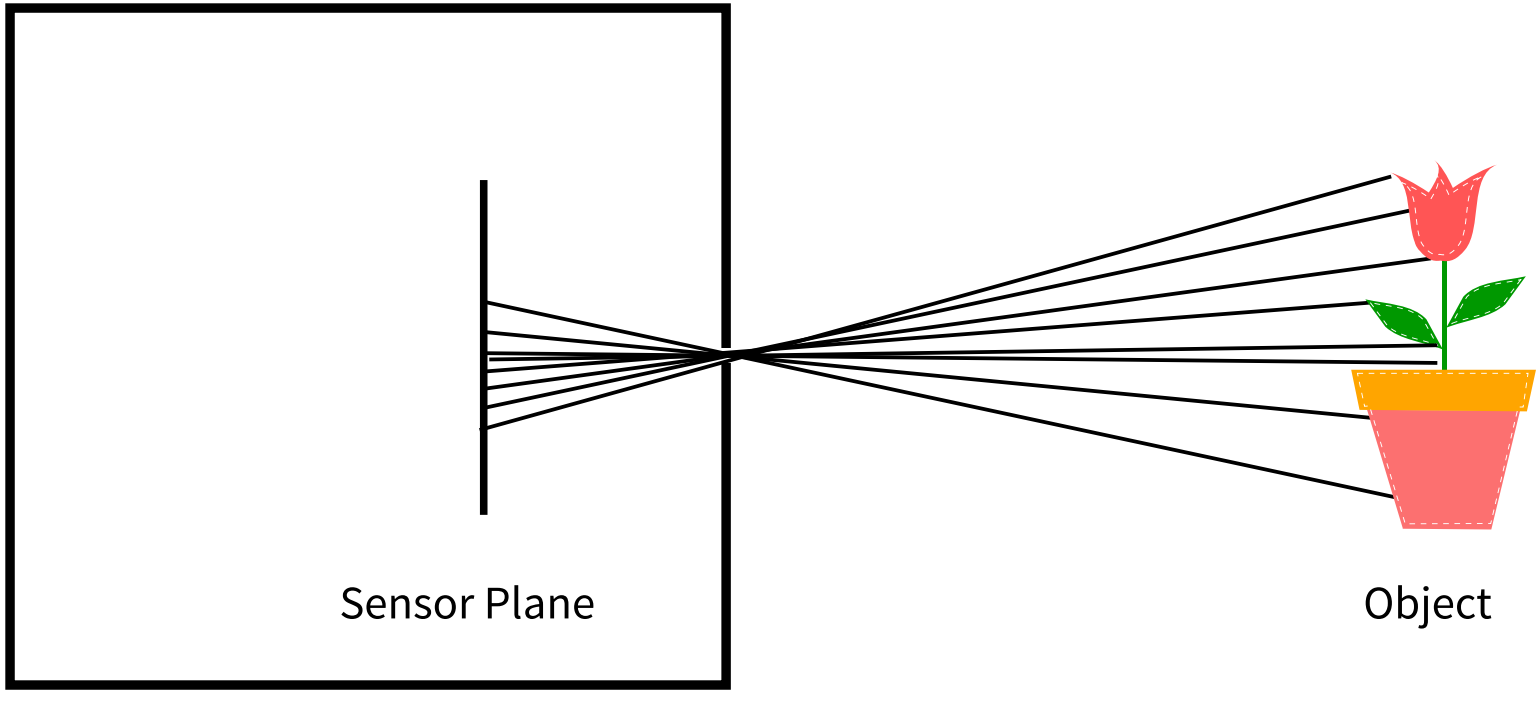
Pinhole Camera

THE PINHOLE CAMERA



Pinhole Camera

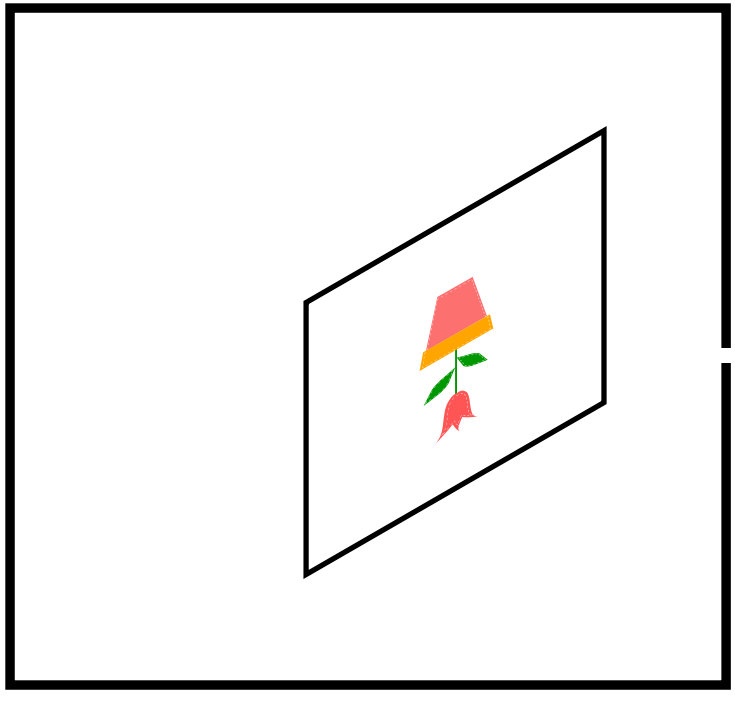
THE PINHOLE CAMERA

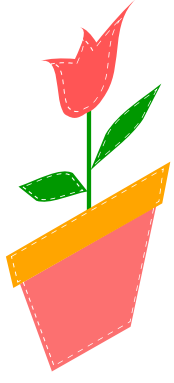


Pinhole Camera

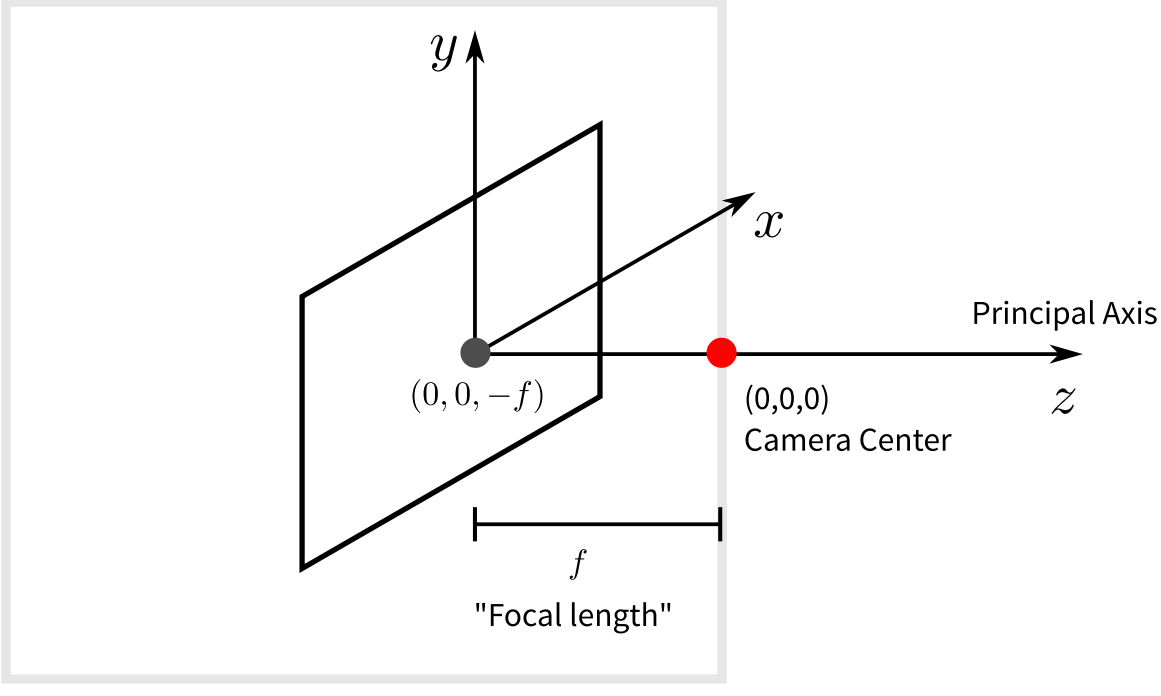
Every point on the sensor plane corresponds to a unique ray

THE PINHOLE CAMERA

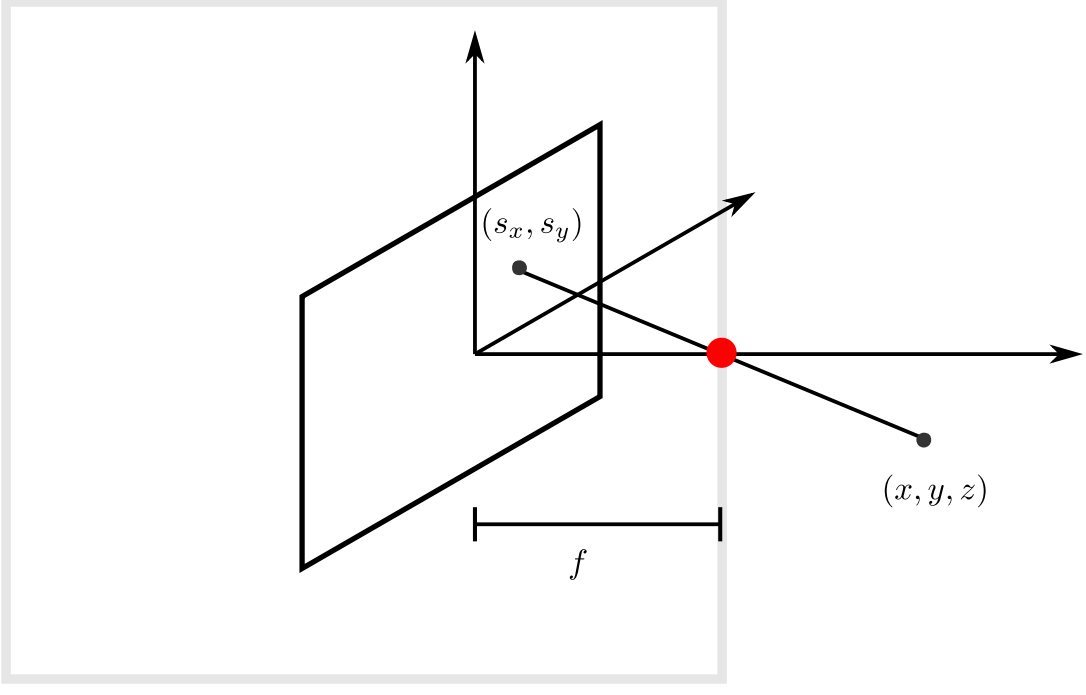




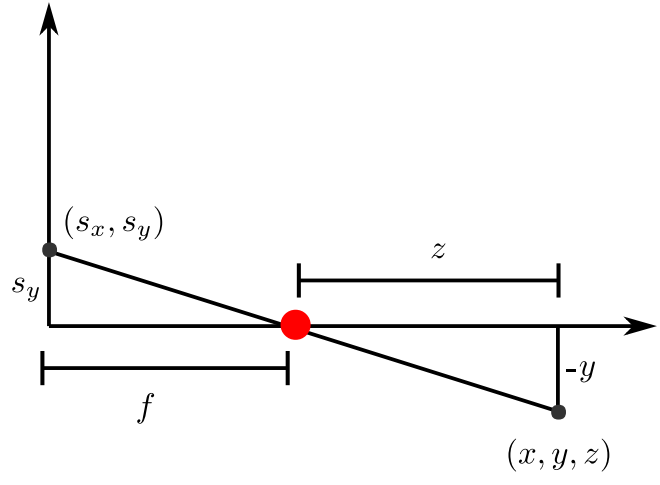
THE PINHOLE CAMERA



THE PINHOLE CAMERA

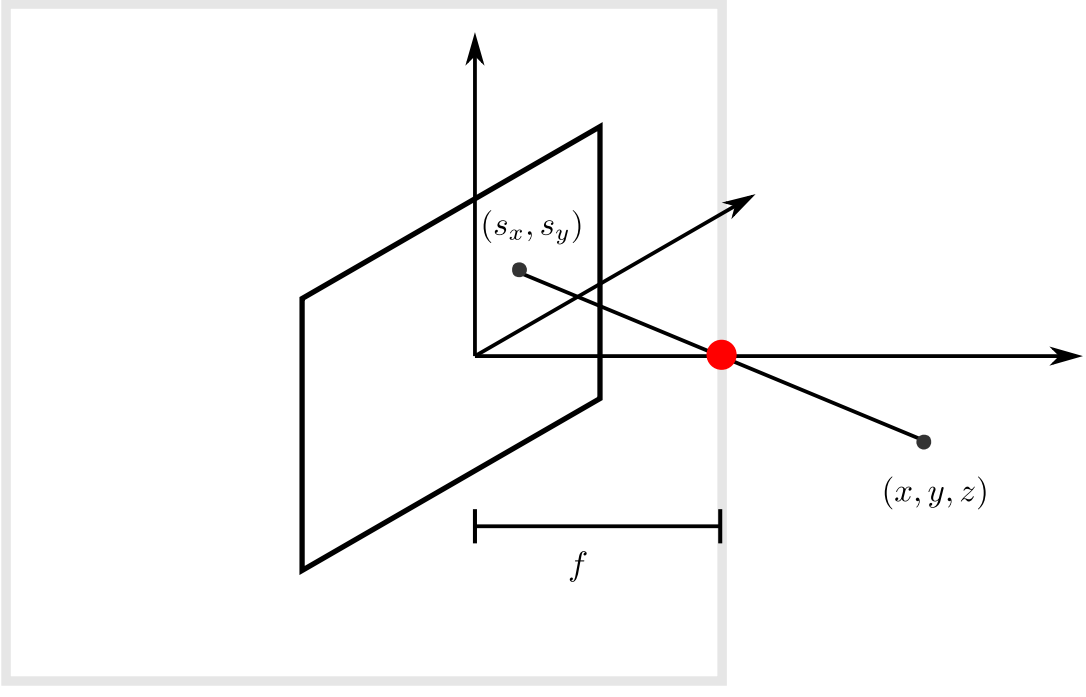


THE PINHOLE CAMERA



By Similar Triangles: $s_y = -f\frac{y}{z}$

THE PINHOLE CAMERA

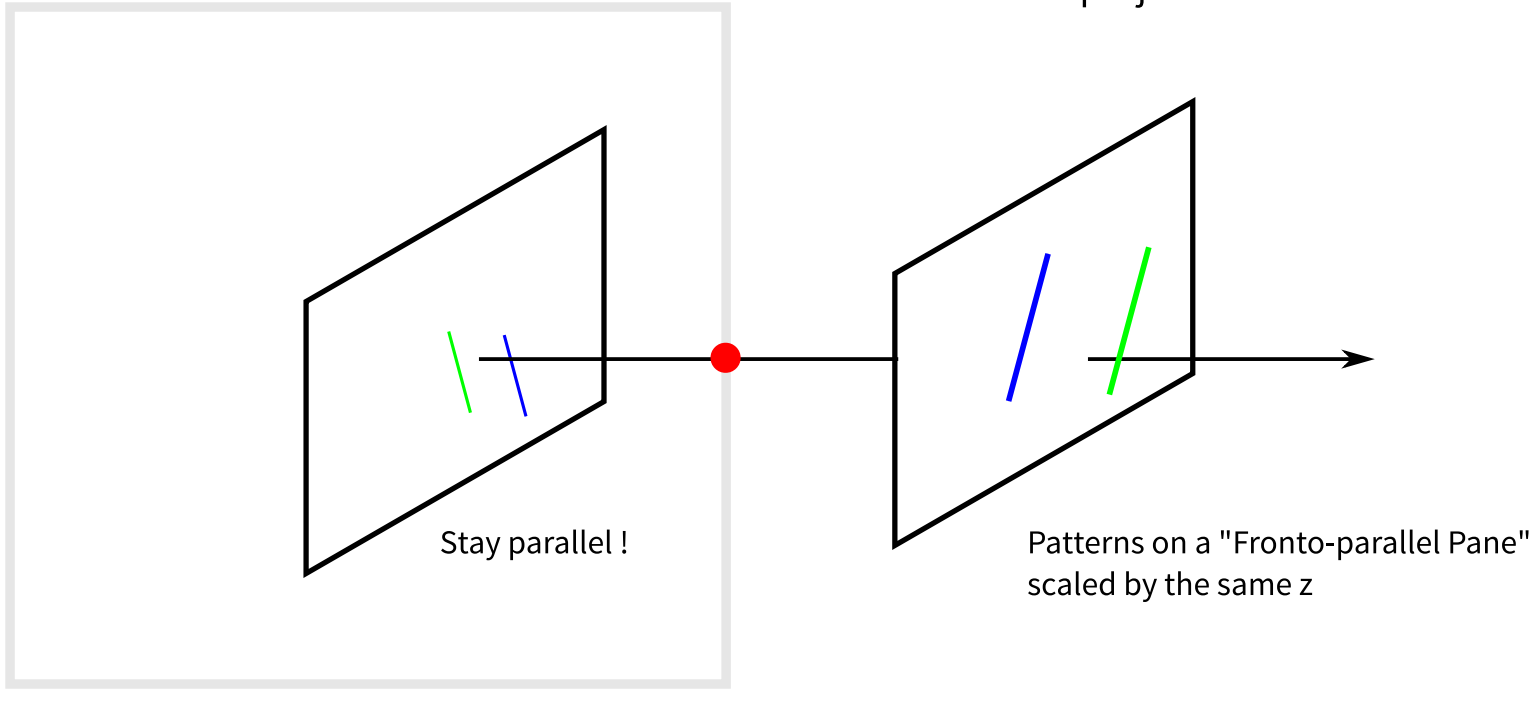


By Similar Triangles: $s_y = -f\frac{y}{z}$ $\qquad$ $s_x = -f\frac{x}{z}$

$$(x, y, z) \Rightarrow \left(-f\frac{x}{z}, -f\frac{y}{z}\right)$$

THE PINHOLE CAMERA

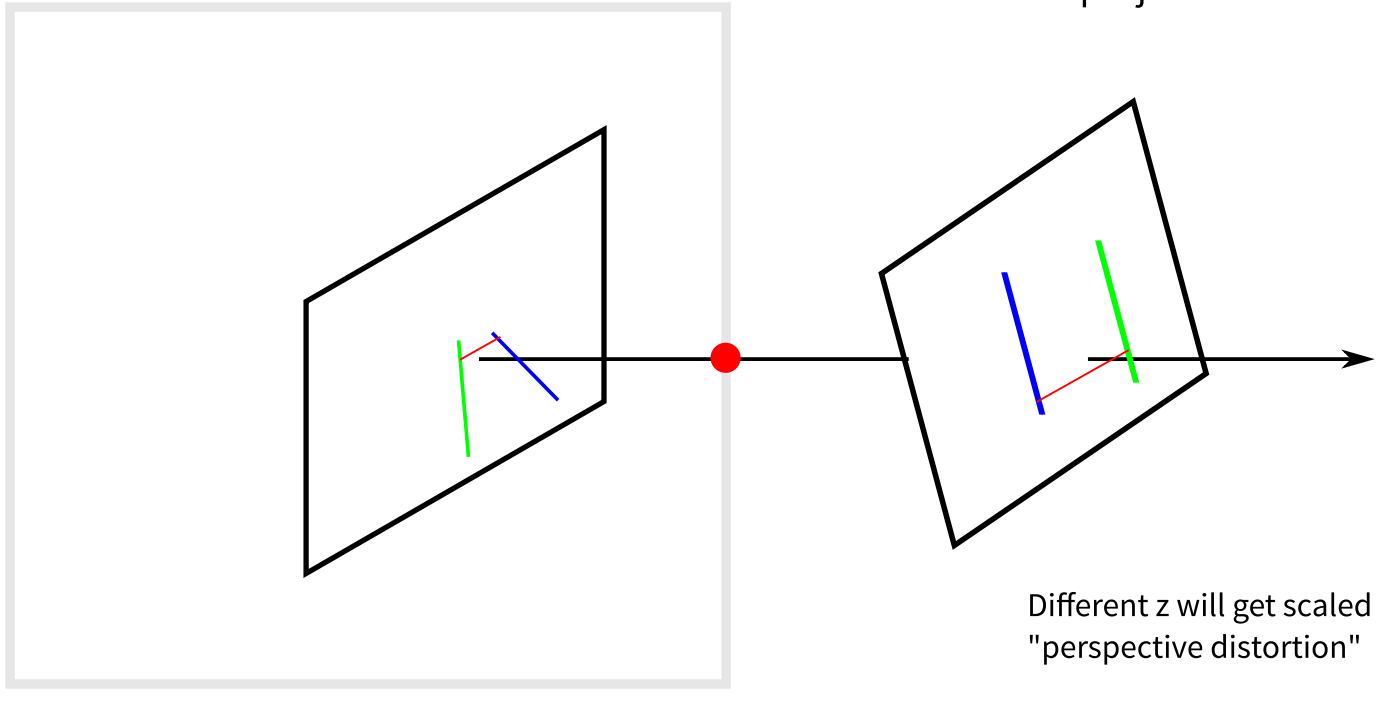
We will go in-depth into camera projection in a few weeks ....



$$(x, y, z) \Rightarrow \left(-f\frac{x}{z}, -f\frac{y}{z}\right)$$

THE PINHOLE CAMERA

We will go in-depth into camera projection in a few weeks ....

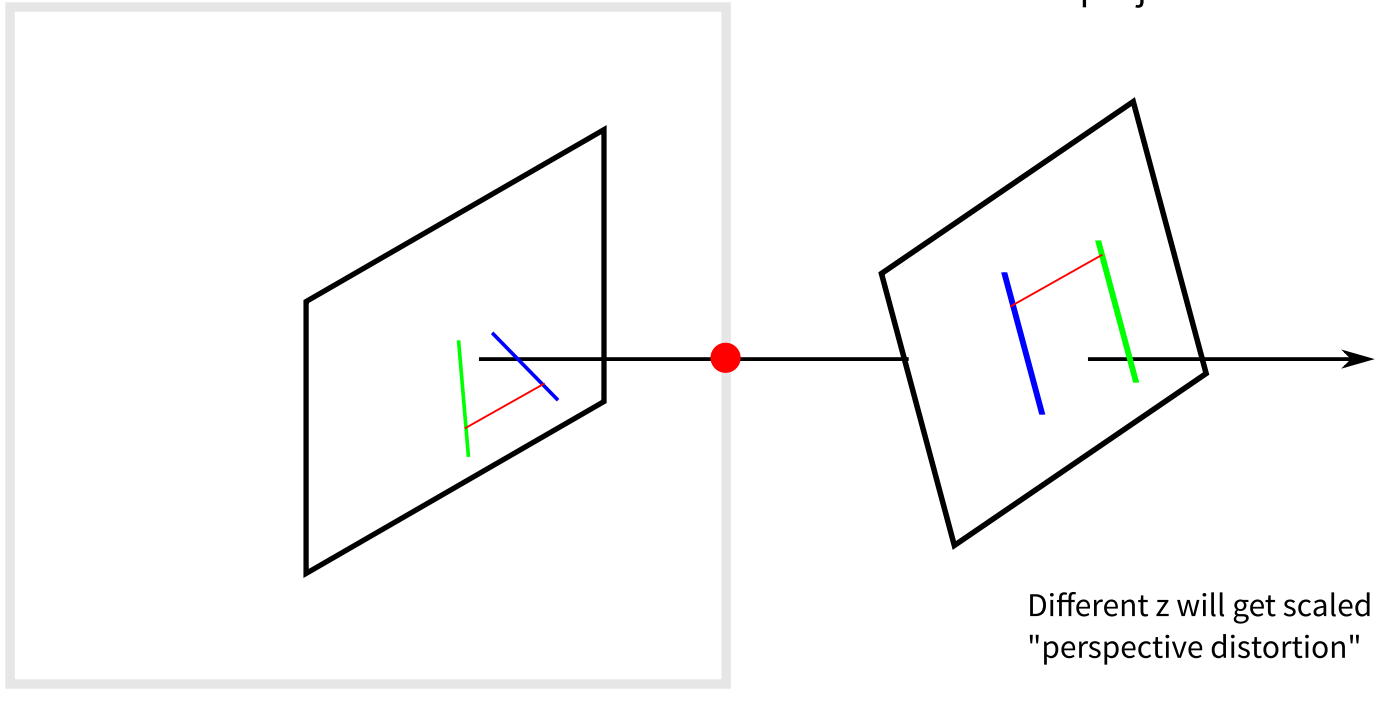


Different z will get scaled differently:
"perspective distortion"

$$(x, y, z) \Rightarrow \left(-f\frac{x}{z}, -f\frac{y}{z}\right)$$

THE PINHOLE CAMERA

We will go in-depth into camera projection in a few weeks ....

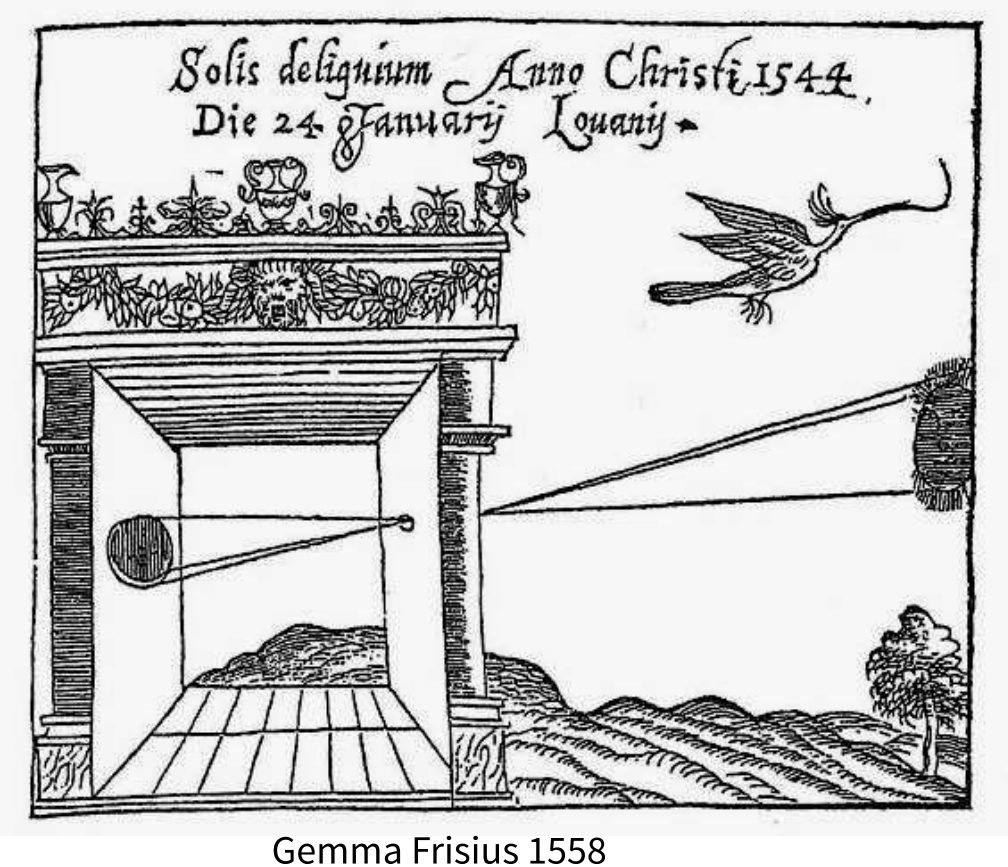


Different z will get scaled differently:
"perspective distortion"

$$(x, y, z) \Rightarrow \left(-f\frac{x}{z}, -f\frac{y}{z}\right)$$

THE PINHOLE CAMERA

Not new to digital cameras!



Gemma Frisius 1558

Basic principle known to Mozi (470-390 BCE), Aristotle (384-322 BCE)

Source: A. Efros

THE PINHOLE CAMERA

Not new to digital cameras!

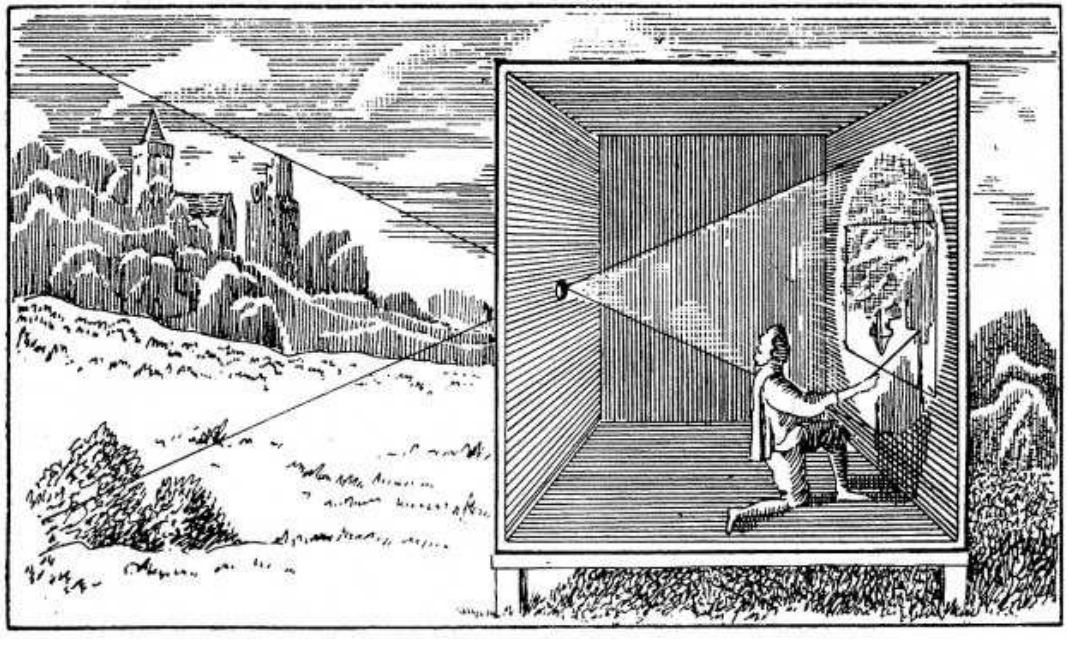


Image via Subhransu Maji

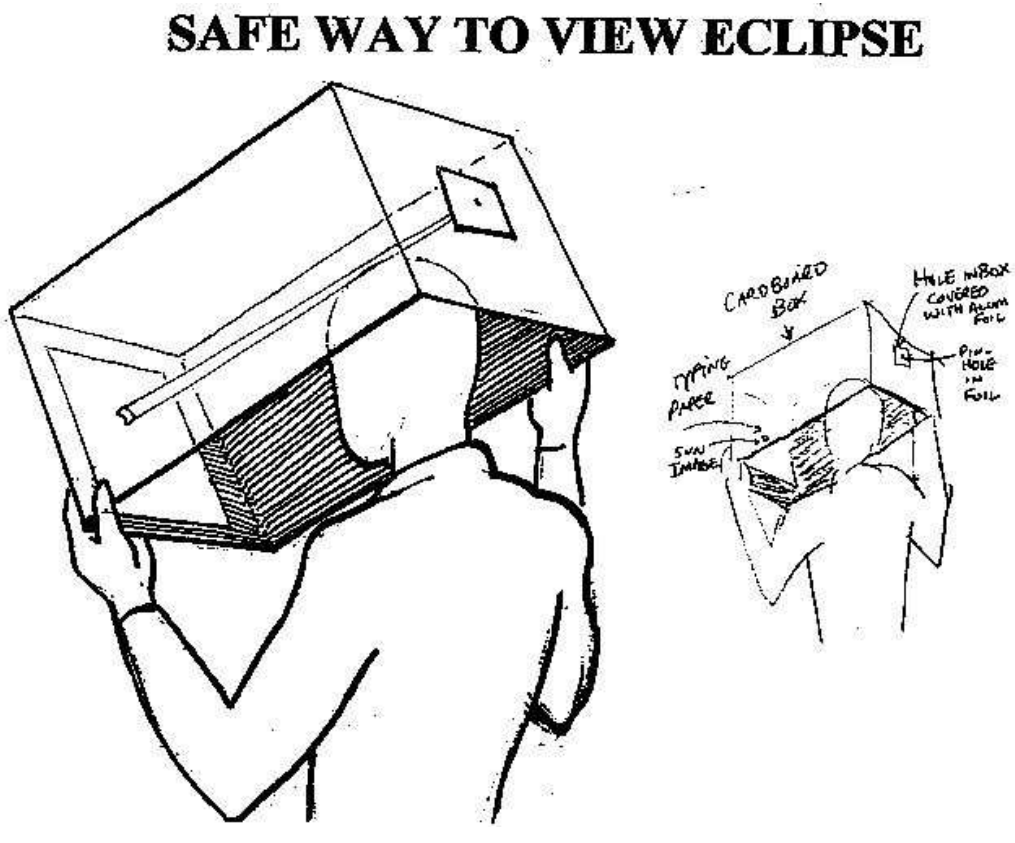
Basic principle known to Mozi (470-390 BCE), Aristotle (384-322 BCE)

Drawing aid for artists: described by Leonardo da Vinci (1452-1519)

Source: A. Efros

THE PINHOLE CAMERA

Eclipsomania 2017 !

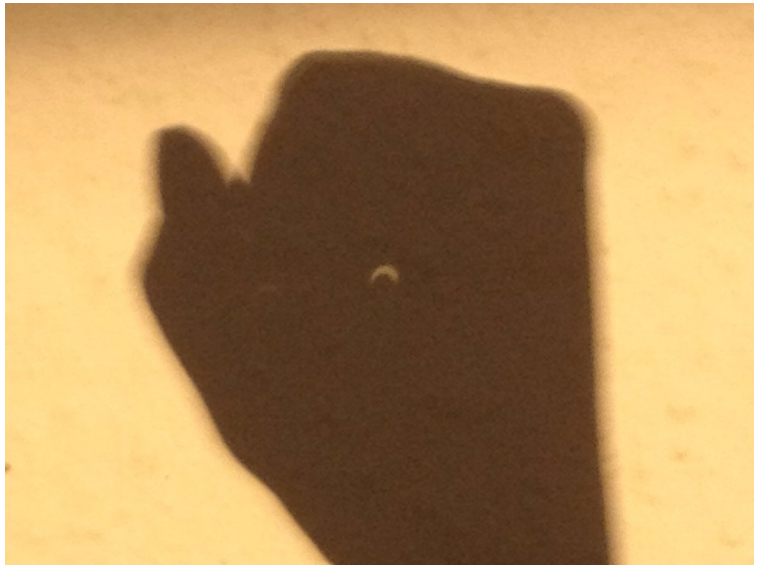


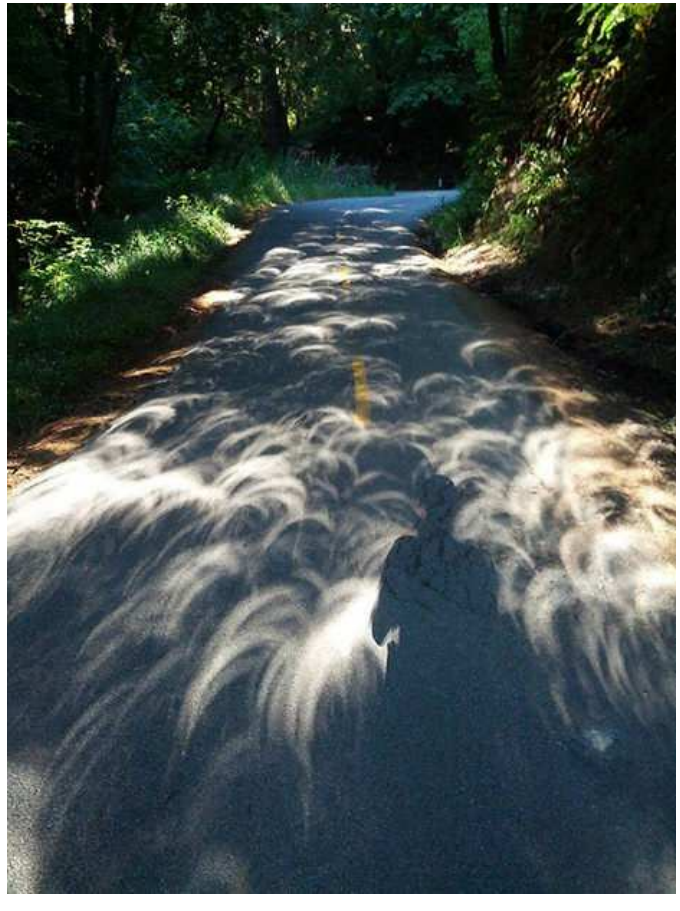
THE PINHOLE CAMERA

Eclipsomania 2017 !

THE PINHOLE CAMERA

Eclipsomania 2017 !

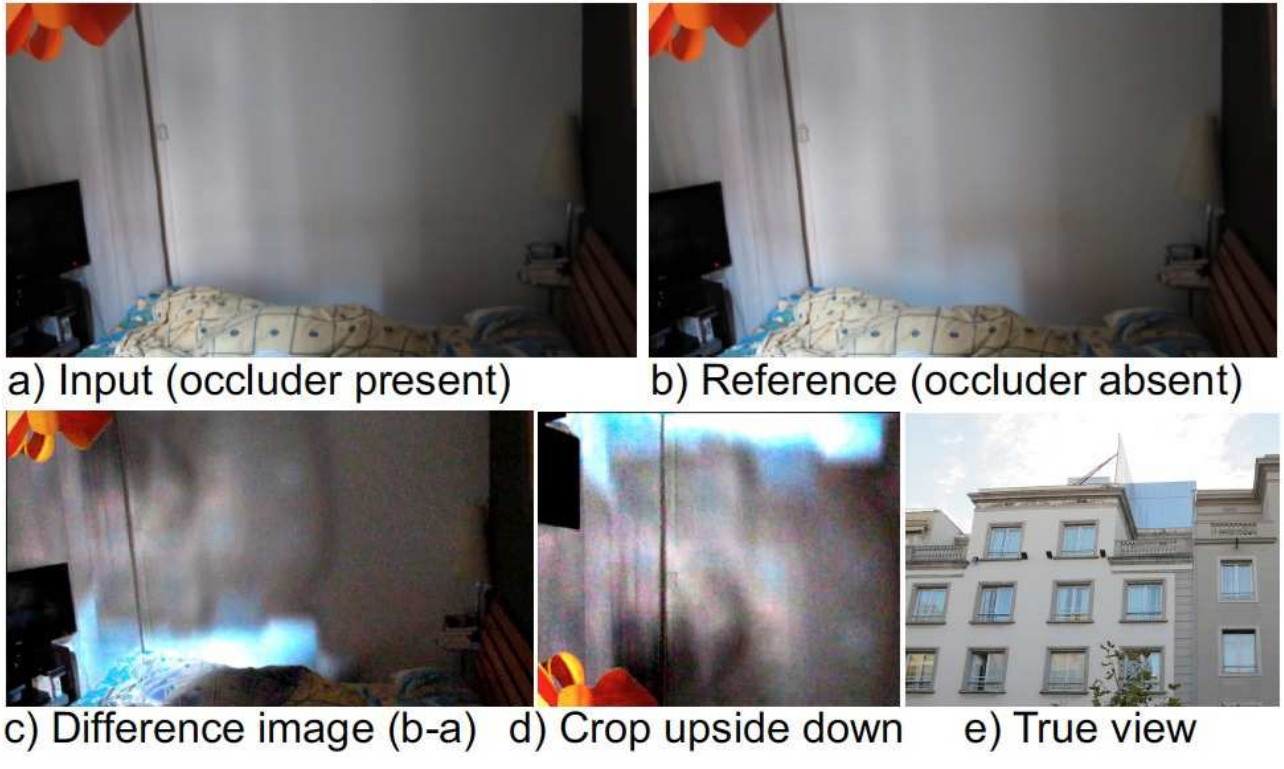




THE PINHOLE CAMERA

Accidental Pinhole and Pinspeck Cameras: Revealing the scene outside the picture
Antonio Torralba, William T. Freeman CVPR 2012
http://people.csail.mit.edu/torralba/research/accidentalcameras/

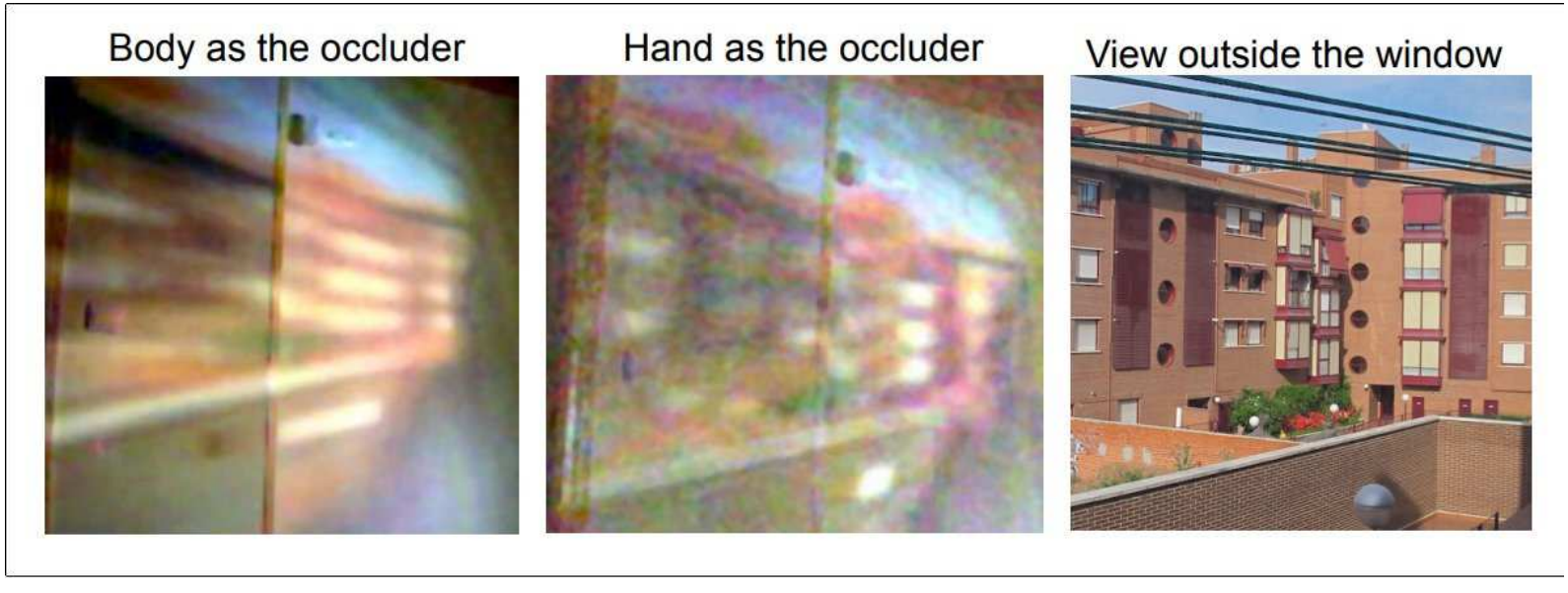
Turning windows and shadows into pinhole and anti-pinhole cameras.



THE PINHOLE CAMERA

Accidental Pinhole and Pinspeck Cameras: Revealing the scene outside the picture
Antonio Torralba, William T. Freeman CVPR 2012
http://people.csail.mit.edu/torralba/research/accidentalcameras/

Turning windows and shadows into pinhole and anti-pinhole cameras.



THE PINHOLE CAMERA

Accidental Pinhole and Pinspeck Cameras: Revealing the scene outside the picture

Antonio Torralba, William T. Freeman CVPR 2012
http://people.csail.mit.edu/torralba/research/accidentalcameras/

Turning windows and shadows into pinhole and anti-pinhole cameras.